

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:31:15 ; Search time 70.555 Seconds  
(without alignments)  
2650.131 Million cell updates/sec

Title: US-09-830-972-29  
Perfect score: 5923  
Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	5882	99.3	1178	21	AAAY71311	Human neurite grow	
2	5815	98.2	1192	22	AAU04591	Human Nogo protein	
3	5815	98.2	1192	23	ABP68600	Human pancreatic c	
4	5810	98.1	1192	21	AAAY56967	Human MAGI polypep	
5	5810	98.1	1192	22	AAB82349	Human NOGO-A prote	
6	5810	98.1	1192	23	ABB81078	Human neurotransmi	
7	5810	98.1	1192	23	ABG30938	Human NogoA protei	
8	5526.5	93.3	1246	22	AAU33228	Novel human secret	
9	4560	77.0	983	24	ABU11573	Human MDDT polypep	
10	4400	74.3	893	21	AAAY95012	Human secreted pro	
11	4296.5	72.5	1163	21	AAAY71310	Rat neurite growth	
12	4296.5	72.5	1163	23	ABB81074	Rat neurotransmitt	
13	4294.5	72.5	1162	21	AAAY71557	Rat Nogo A truncat	
14	4286.5	72.4	1163	21	AAAY71384	Alternative versio	
15	3388.5	57.2	974	21	AAAY71560	Rat Nogo A protein	
16	3146.5	53.1	642	19	AAW58383	Human secreted pro	
17	3146.5	53.1	642	22	AAB90682	Human BG160_1 prot	
18	2715	45.8	803	21	AAAY71562	Rat Nogo A protein	
19	2529.5	42.7	737	21	AAAY71386	Rat Nogo A protein	
20	2487.5	42.0	746	21	AAAY71391	Rat Nogo A protein	
21	2457	41.5	736	21	AAAY71398	Rat Nogo A protein	
22	2449.5	41.4	732	21	AAAY71399	Rat Nogo A protein	
23	2405.5	40.6	695	21	AAAY71387	Rat Nogo A protein	
24	2344.5	39.6	684	21	AAAY71394	Rat Nogo A protein	
25	1948.5	32.9	552	21	AAAY71388	Rat Nogo A protein	
26	1743	29.4	502	21	AAAY71396	Rat Nogo A protein	
27	1634.5	27.6	475	21	AAAY71389	Rat Nogo A protein	
28	1566.5	26.4	403	21	AAAY71563	Rat Nogo A protein	
29	1552.5	26.2	457	21	AAAY71392	Rat Nogo A protein	
30	1495.5	25.2	373	21	AAB24242	Human Nogo B prote	
31	1495.5	25.2	373	21	AAAY56969	Human MAGI polypep	
32	1495.5	25.2	373	21	AAAY53624	A bone marrow secr	
33	1495.5	25.2	373	22	AAB82350	Human NOGO-B prote	
34	1495.5	25.2	373	23	ABP68601	Human pancreatic c	
35	1495.5	25.2	373	23	ABB81079	Human neurotransmi	
36	1495.5	25.2	373	23	AAM47954	Human RTN4B SEQ ID	
37	1487.5	25.1	373	23	ABG30937	Human NogoB protei	
38	1412	23.8	289	21	AAAY56968	Human MAGI polypep	
39	1327	22.4	284	21	AAAY95030	Human clone vb22_1	
40	1264	21.3	356	21	AAAY71390	Rat Nogo A protein	
41	1205.5	20.4	374	21	AAAY71397	Rat Nogo A protein	
42	1196.5	20.2	361	21	AAAY71385	Alternative versio	
43	1189	20.1	359	21	AAAY71558	Rat Nogo A protein	
44	1187	20.0	360	21	AAAY71383	Rat neurite growth	
45	1187	20.0	360	23	ABB81076	Rat neurotransmitt	

# ALIGNMENTS

RESULT 1

AAY71311

ID AAY71311 standard; Protein; 1178 AA.

XX

AC AAY71311;

XX

DT 02-NOV-2000 (first entry)

XX

DE Human neurite growth inhibitor Nogo.

XX

KW Human; neurite growth inhibitor; Nogo; neural cell; myelin; CNS;  
 KW central nervous system; neoplastic disease; antiproliferative; glioma;  
 KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
 KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
 KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
 KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
 KW structural plasticity; screening.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 187

FT /label= Unknown

FT Misc-difference 188

FT /label= Unknown

FT Misc-difference 189

FT /label= Unknown

FT Misc-difference 190

FT /label= Unknown

FT Misc-difference 221

FT /label= Unknown

FT Misc-difference 328

FT /label= Unknown

FT Misc-difference 477

FT /label= Unknown

FT Region 994..1174

FT /note= "Region specifically described in claim 16"

FT Region 977..1012

FT /note= "Region specifically described in claim 16"

FT Region 1079..1114

FT /note= "Region specifically described in claim 16"

XX

PN WO200031235-A2.

XX

PD 02-JUN-2000.

XX

PF 05-NOV-1999; 99WO-US26160.

XX

PR 06-NOV-1998; 98US-0107446.

XX

PA (SCHW/) SCHWAB M E.

PA (CHEN/) CHEN M S.

XX

PI Schwab ME, Chen MS;

XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic  
PT disorders of the central nervous system and inducing regeneration of  
PT neurons -

XX

PS Claim 11; Fig 13; 122pp; English.

XX

CC The present sequence is a human Nogo protein which is a  
CC potent neural cell growth inhibitor and is free of all central nervous  
CC system (CNS) myelin material with which it is natively associated. The  
CC human Nogo sequence was derived by aligning human expressed sequence tags  
CC (ESTs) e.g. AA158636, AA333267, AA081783, AA167765, AA322918, AA092565,  
CC AA081525 and AA081840 with the rat Nogo sequence.  
CC Nogo proteins and fragments displaying neurite growth inhibitory  
CC activity are used in the treatment of neoplastic disease of the CNS  
CC e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma, ependyoma,  
CC pinealoma, haemangioblastoma, acoustic neuroma, oligodendroglioma,  
CC meningioma, neuroblastoma or retinoblastoma and degenerative nerve  
CC diseases e.g. Alzheimer's and Parkinson's diseases. Therapeutics which  
CC promote Nogo activity can be used to treat or prevent hyperproliferative  
CC or benign dysproliferative disorders e.g. psoriasis and tissue  
CC hypertrophy. Ribozymes or antisense Nogo nucleic acids can be used to  
CC inhibit production of Nogo protein to induce regeneration of neurons or  
CC to promote structural plasticity of the CNS in disorders where neurite  
CC growth, regeneration or maintenance are deficient or desired.  
CC The animal models can be used in diagnostic and screening methods for  
CC predisposition to disorders and to screen for or test molecules which  
CC can treat or prevent disorders or diseases of the CNS.  
CC Note: SEQ ID numbers 35-42 are referred in claim 32 and SEQ ID NO: 29  
CC in disclosure of the specification. However the specification does not  
CC include sequences for these SEQ ID numbers.

XX

SQ Sequence 1178 AA;

Query Match 99.3%; Score 5882; DB 21; Length 1178;  
Best Local Similarity 99.7%; Pred. No. 1.8e-297;  
Matches 1175; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Db	1	MEDLDQSPLVSSSDSVPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
QY	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
QY	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKRRG	180
Db	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKRRG	180
QY	181	SSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAAFKEHEYLGN	240
Db	181	SSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAAFKEHEYLGN	240
QY	241	LSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFSVSPKAESAV	300
Db	241	LSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFSVSPKAESAV	300



Qy	301	IVANPREEII VKNKDEEEKLVSNNILHXQQELPTALT KL VKED EVVSSEKAKDSFNEKRV	360
Db	301	IVANPREEII VKNKDEEEKLVSNNILHXQQELPTALT KL VKED EVVSSEKAKDSFNEKRV	360
Qy	361	AVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTNH	420
Db	361	AVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTNH	420
Qy	421	EKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPILLEDPTSENXTDE	480
Db	421	EKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPILLEDPTSENXTDE	480
Qy	481	KKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLT KVTEE VVANMPEGLTPDL	540
Db	481	KKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLT KVTEE VVANMPEGLTPDL	540
Qy	541	VQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDI	600
Db	541	VQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDI	600
Qy	601	VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIKE	660
Db	601	VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIKE	660
Qy	661	EIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELV	720
Db	661	EIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELV	720
Qy	721	EDSSPDSEPVDLFSDDSI PDVPQKQDET VMLVKESLTETSFESMIEYENKEKLSALPPEG	780
Db	721	EDSSPDSEPVDLFSDDSI PDVPQKQDET VMLVKESLTETSFESMIEYENKEKLSALPPEG	780
Qy	781	GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIP LQMEELSTAVYSNDDL FISKEAQIRE	840
Db	781	GKPYLESFKLSLVNTKDTLLPDEVSTLSKKEKIP LQMEELSTAVYSNDDL FISKEAQIRE	840
Qy	841	TETFSDDSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL	900
Db	841	TETFSDDSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL	900
Qy	901	PHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDV SALGHTQAEIESIVKPKVLE	960
Db	901	PHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDV SALGHTQAEIESIVKPKVLE	960
Qy	961	KEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDI KKTGVVFGASLFLLLSLTVF	1020
Db	961	KEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDI KKTGVVFGASLFLLLSLTVF	1020
Qy	1021	SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS	1080
Db	1021	SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS	1080
Qy	1081	ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI	1140
Db	1081	ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI	1140
Qy	1141	YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178

Db 1141 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178

RESULT 2

AAU04591

ID AAU04591 standard; Protein; 1192 AA.

XX

AC AAU04591;

XX

DT 26-SEP-2001 (first entry)

XX

DE Human Nogo protein.

XX

KW Human; Nogo receptor; axonal growth; immunogen; antibody; nogo protein;  
KW cranial trauma; cerebral trauma; spinal cord injury; stroke;  
KW demyelinating disease; multiple sclerosis; monophasis demyelination;  
KW encephalomyelitis; multifocal leukoencephalopathy; panencephalitis;  
KW Marchiafava-Bignami disease; pontine myelinolysis; adrenoleukodystrophy;  
KW Pelizaeus-Merzbacher disease; Spongy degeneration; Alexander's disease;  
KW Canavan's disease; metachromatic leukodystrophy; viral infection;  
KW Krabbe's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain

1054..1119

FT /label= Lumenal\_extracellular\_domain

FT /note= "This sequence is specifically claimed"

FT Peptide

1055..1094 ✓

FT /label= Pep1

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide

1064..1088

FT /label= Pep2

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide

1074..1098

FT /label= Pep3

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide

1084..1108

FT /label= Pep4

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically cly

FT Peptide

1095..1119

FT /label= Pep5

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

XX

PN WO200151520-A2.

XX

PD 19-JUL-2001.

XX

PF 12-JAN-2001; 2001WO-US01041.

XX

PR 12-JAN-2000; 2000US-0175707.

-- 18550X-

PR 26-MAY-2000; 2000US-0207366.  
PR 29-SEP-2000; 2000US-0236378.

XX

PA (UYYA ) UNIV YALE.

XX

PI Strittmatter SM;

XX

DR WPI; 2001-442138/47.

DR N-PSDB; AAS09453.

XX

PT Novel Nogo receptor protein useful for identifying modulator of Nogo  
PT protein or Nogo receptor protein, which is useful for treating central  
PT nervous system disorders -

XX

PS Example 1; Page 101-104; 109pp; English.

XX

CC The sequence is the human Nogo protein, a 250kDa myelin-associated axon  
CC growth inhibitor. The invention relates to the use of the nogo receptor,  
CC nogo protein, their nucleic acids, vectors expressing them and antibodies  
CC against them, to isolate agents which block nogo receptor mediated axonal  
CC growth. The agent is useful for treating a central nervous system  
CC disorder which is a result of cranial or cerebral uma, spinal cord  
CC injury, stroke or a demyelinating disease selected from multiple  
CC sclerosis, monophasis demyelination, encephalomyelitis, multifocal  
CC leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease,  
CC pontine myelinolysis, adrenoleukodystrophy, Pelizaeus-Merzbacher disease,  
CC Spongy degeneration, Alexander's disease, Canavan's disease,  
CC metachromatic leukodystrophy, viral infection and Krabbe's disease.

XX

SQ Sequence 1192 AA;

Query Match 98.2%; Score 5815; DB 22; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 5.6e-294;  
Matches 1160; Conservative 4; Mismatches 13; Indels 16; Gaps 3;

```
QY      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPMVSXSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLE  A 60

QY     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

QY    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

QY    181 SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:
Db    181 SSGSVDETLFALPAASEFPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

QY    227 LSAASFKEHEYLGNLSTVLPTTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    241 LSAASFKEHEYLGNLSTVLPTTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

QY    287 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV 346
      |||
```

Db 301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNNQELPTALTCLKVKEDEVV 360  
 Qy 347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406  
 |||||  
 Db 361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420  
 Qy 407 DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466  
 |||||  
 Db 421 DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480  
 Qy 467 PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTQVTE 526  
 |||||  
 Db 481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTQVTE 540  
 Qy 527 EVVANMPEGLTPDLVQEACESELNEVTGKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586  
 |||||  
 Db 541 EVVANMPEGLTPDLVQEACESELNEVTGKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600  
 Qy 587 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 646  
 |||||  
 Db 601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 660  
 Qy 647 EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPPDFSDYSE 705  
 |||||  
 Db 661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPPDFSDYSE 720  
 Qy 706 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI 765  
 |||||  
 Db 721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI 780  
 Qy 766 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTA  
 |||||  
 Db 781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTAVY 840  
 Qy 826 SNDDLFI SKEAQIRETETFS DSSPIEIIDEFPTLISSKTD SFSKLAREYTDLEVSHKSEI 885  
 |||||  
 Db 841 SNDDLFI SKEAQIRETETFS DSSPIEIIDEFPTLISSKTD SFSKLAREYTDLEVSHKSEI 900  
 Qy 886 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALGH 945  
 |||||  
 Db 901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALA- 959  
 Qy 946 TQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV 1005  
 |||||  
 Db 960 TQAEIESIVKPKVLVKEAEKKLP SDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV 1019  
 Qy 1006 VFGASLFLLLSLTVFSIVSVTAYIALALLSVT SFRIYKGVIAIQKSDEGHPFRAYLES 1065  
 |||||  
 Db 1020 VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES 1079  
 Qy 1066 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT 1125  
 |||||  
 Db 1080 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT 1139  
 Qy 1126 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 |||||  
 Db 1140 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 3

ABP68600

ID ABP68600 standard; Protein; 1192 AA.

XX

AC ABP68600;

XX

DT 14-JAN-2003 (first entry)

XX

DE Human pancreatic cancer expressed protein SEQ ID NO 71.

XX

KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;  
KW cytostatic; tumour.

XX

OS Homo sapiens.

XX

PN WO200260317-A2.

XX

PD 08-AUG-2002.

XX

PF 30-JAN-2002; 2002WO-US02781.

XX

PR 30-JAN-2001; 2001US-265305P.

PR 31-JAN-2001; 2001US-265682P.

PR 09-FEB-2001; 2001US-267568P.

PR 21-MAR-2001; 2001US-278651P.

PR 28-APR-2001; 2001US-287112P.

PR 16-MAY-2001; 2001US-291631P.

PR 12-JUL-2001; 2001US-305484P.

PR 20-AUG-2001; 2001US-313999P.

PR 27-NOV-2001; 2001US-333626P.

XX

PA (CORI-) CORIXA CORP.

XX

PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;

XX

DR WPI; 2002-627435/67.

DR N-PSDB; ABV94680.

XX

PT New isolated polynucleotide and pancreatic tumor polypeptides, useful  
PT for diagnosing, preventing and/or treating cancer, particularly  
PT pancreatic cancer -

XX

PS Claim 2; SEQ ID NO 71; 300pp + Sequence Listing; English.

XX

CC The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);  
CC (b) complements of (a); (c) sequences consisting of at least 20  
CC contiguous residues of (a); (d) sequences that hybridize to (a), under  
CC moderately stringent conditions; (e) sequences having at least 75% or 90%  
CC identity to (a); or (f) degenerate variants of (a). Polypeptides  
CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to  
CC detect cancer in a patient and compositions comprising polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations and  
CC antigen presenting cells expressing the polypeptide are useful in  
CC treating pancreatic cancer and stimulating an immune response. The

CC polynucleotides can be used as probes or primers for nucleic acid  
CC hybridisation, in the design and preparation of ribozyme molecules for  
CC inhibiting expression of the tumour polypeptides and proteins in the  
CC tumour cells, in vaccines and for gene therapy.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 1192 AA;

Query Match 98.2%; Score 5815; DB 23; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 5.6e-294;  
Matches 1160; Conservative 4; Mismatches 13; Indels 16; Gaps 3;

```
Qy      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP 120

Qy    121 SPLSAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:
Db    181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy    227 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy    287 GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV 346
      |||
Db    301 GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDEVV 360

Qy    347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406
      |||
Db    361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy    407 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466
      |||
Db    421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy    467 PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 526
      |||
Db    481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 540

Qy    527 EVVANMPEGLTPDLVQEACESELNEVTGTGIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586
      |||
Db    541 EVVANMPEGLTPDLVQEACESELNEVTGTGIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy    587 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 646
      |||
Db    601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 660
```



```

PD      03-FEB-2000 .
XX
PF      21-JUL-1999;      99WO-GB02360 .
XX
PR      22-JUL-1998;      98GB-0016024 .
PR      19-JUL-1999;      99GB-0016898 .
XX
PA      (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI      Michalovich D,   Prinjha RK;
XX
DR      WPI; 2000-182693/16.
DR      N-PSDB; AAZ56886.

```

Query Match 98.1%; Score 5810; DB 21; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 1e-293;  
Matches 1159; Conservative 4; Mismatches 14; Indels 16; Gaps 3;

Qy	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLLEEVLERKPA	60
Db	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLLEEVLERKPA	60
Qy	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDFSPVSSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDFSPVSSTVPAP	120
Qy	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Db	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Qy	181	SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP	226
Db	181	SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP	240
Qy	227	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	286
Db	241	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	300
Qy	287	GSSFSVSPKAESAVIVANPREEIIIVKNKDEEEKLVSNNIILHXQQELPTALTCLKVKEDEVV	346



Db	301		GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDEVV	360
Qy	347		SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	406
Db	361		SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	420
Qy	407		DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF	466
Db	421		DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF	480
Qy	467		PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVTE	526
Db	481		PLLGDPPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE	540
Qy	527		EVVANMP TPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSP	586
Db	541		EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSP	600
Qy	587		EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601		EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647		EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFS DYSE	705
Db	661		EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFS DYSE	720
Qy	706		MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	765
Db	721		MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	780
Qy	766		EYENKEKLSALPPEGCKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	825
Db	781		EYENKEKLSALPPEGCKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	840
Qy	826		SNDDLFI SKEAQIRETETFS DSSPIEIIDEFPTLISSKTD SFSKLAREYTDLEVSHKSEI	885
Db	841		SNDDLFI SKEAQIRETETFS DSSPIEIIDEFPTLISSKTD SFSKLAREYTDLEVSHKSEI	900
Qy	886		ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVVLLPPDV SALGH	945
Db	901		ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVVLLPPDV SALA-	959
Qy	946		TQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGV	1005
Db	960		TQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFS AE SVVDLLYWRDIKKTGV	1019
Qy	1006		VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1065
Db	1020		VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1079
Qy	1066		EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1125
Db	1080		EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1139
Qy	1126		LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178





Db	721	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	840
Qy	826	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLI SSKTDSFSKLAREYTDLEVSHKSEI	885
Db	841	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLI SSKTDSFSKLAREYTDLEVSHKSEI	900
Qy	886	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI SFSDDFSKNGSATS KVL LPPDV SALGH	945
Db	901	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI SFSDDFSKNGSATS KVL LPPDV SALA-	959
Qy	946	TQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAI FSADLGKTSVVDLLYWRDI KKTGV	1005
Db	960	TQAEIESIVKPKVLVKEAEKKLP SDTEKEDRSPSAI FSAELSKTSVVDLLYWRDI KKTGV	1019
Qy	1006	VFGASLFLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES	1065
Db	1020	VFGASLFLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES	1079
Qy	1066	EVAISEELVQKYSNSALGHVNCTI KELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1125
Db	1080	EVAISEELVQKYSNSALGHVNCTI KELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1139
Qy	1126	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA KIPGLKRKAE	1178
Db	1140	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA KIPGLKRKAE	1192

RESULT 6

ABB81078

ID ABB81078 standard; Protein; 1192 AA.

XX

AC ABB81078;

XX

DT 05-NOV-2002 (first entry)

XX

DE Human neurotransmitter receptor protein Nogo-A.

XX

KW Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;  
 KW central nervous system; peripheral nervous system; tranquilizer; Nogo;  
 KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;  
 KW nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;  
 KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;  
 KW neurotransmitter receptor; human; receptor.

XX

OS Homo sapiens.

XX

PN US2002072493-A1.

XX

PD 13-JUN-2002.

XX

PF 28-JUN-2001; 2001US-0893348.

XX

PR 19-MAY-1998; 98IL-0124500.  
PR 21-JUL-1998; 98WO-US14715.  
PR 22-DEC-1998; 98US-0218277.  
PR 19-MAY-1999; 99US-0314161.

XX

PA (YEDA ) YEDA RES & DEV CO LTD.

XX

PI Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonogo A;  
PI Moalem G;

XX

DR WPI; 2002-607255/65.

DR N-PSDB; ABN86601.

XX

PT Promoting nerve regeneration and preventing neuronal degeneration in  
PT the central/peripheral nervous system from injury/disease, comprises  
PT administering nervous system-specific activated T cells/antigen, or  
PT analogs/peptides -

XX

PS Examples; Page 53-56; 93pp; English.

XX

CC The invention relates to promoting nerve regeneration or conferring  
CC neuroprotection and preventing or inhibiting neuronal degeneration in the  
CC central/peripheral nervous system (NS). The method involves administering  
CC NS-specific activated T cells, NS-specific antigen, its analogue or its  
CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or  
CC combinations. The method is useful for promoting nerve regeneration and  
CC preventing neuronal degeneration in central/peripheral nervous system  
CC from injury/disease, where the injury is spinal cord injury, blunt  
CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or  
CC damages caused by surgery such as tumour excision. The disease is not an  
CC autoimmune disease or neoplasm. The disease results in a degenerative  
CC process occurring in either gray or white matter or both. The disease  
CC is diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's  
CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,  
CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and  
CC vitamin deficiency, intervertebral disc herniation, prion diseases such  
CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral  
CC neuropathies associated with various diseases, including but not limited  
CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute  
CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary  
CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption  
CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-  
CC pathies, complications of various drugs (e.g., metronidazole) and toxins  
CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia  
CC telangiectasia, Friedreich's ataxia, amyloid polyneuropathies,  
CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's  
CC disease, or lipoproteinemia. The present sequence represents the human  
CC neurotransmitter receptor protein Nogo-A, an example of NS-specific  
CC antigen.

XX

SQ Sequence 1192 AA;

Query Match 98.1%; Score 5810; DB 23; Length 1192;

Best Local Similarity 97.2%; Pred. No. 1e-293;

Matches 1159; Conservative 4; Mismatches 14; Indels 16; Gaps 3;

QY 1 MEDLDQSPLVSSSDSPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Db	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Qy	61	AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
Qy	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Db	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Qy	181	SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP	226
Db	181	SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP	240
Qy	227	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	286
Db	241	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	300
Qy	287	GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV	346
Db	301	GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNNQELPTALTCLKVKEDEVV	360
Qy	347	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	406
Db	361	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	420
Qy	407	DKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF	466
Db	421	DKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF	480
Qy	467	PLLEDPTSSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTKVTE	526
Db	481	PLLGDPTSSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVLVAAQDSETDYVTTDNLTKVTE	540
Qy	527	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSP	586
Db	541	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSP	600
Qy	587	ESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601	ESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSE	705
Db	661	EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSE	720
Qy	706	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	765
Db	721	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	840
Qy	826	SNDDLFI SKEAQIRETETFS DSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEI	885

Db 841 SNDDLFISKEAQIRETETFSDDSSPIEIIDEFPTLISSKTDSPSKLAREYTDLEVSHKSEI 900  
 Qy 886 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSQVLLLLPPDVSA LGH 945  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSQVLLLLPPDVSA LA- 959  
 Qy 946 TQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV 1005  
 ||||||||||||| |||||||||||||||||||||:| |||||||||||||||||  
 Db 960 TQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV 1019  
 Qy 1006 VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES 1065  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1020 VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES 1079  
 Qy 1066 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT 1125  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1080 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT 1139  
 Qy 1126 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA KIPGLKRKAE 1178  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1140 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA KIPGLKRKAE 1192

RESULT 7

ABG30938

ID ABG30938 standard; Protein; 1192 AA.

XX

AC ABG30938;

XX

DT 21-OCT-2002 (first entry)

XX

DE Human NogoA protein.

XX

KW Human; Nogo; BACE; acute neuronal injury; spinal injury; head injury;  
 KW stroke; peripheral nerve damage; neoplastic disorder; glioblastoma;  
 KW neuroblastoma; hyperproliferative disorder; dysproliferative disorder;  
 KW cirrhosis; psoriasis; keloid formation; fibrocystic condition; cancer;  
 KW tissue hypertrophy; central nervous system; axon regeneration; NogoA;  
 KW Nogo-associated disease; metastasis.

XX

OS Homo sapiens.

XX

PN WO200257483-A2.

XX

PD 25-JUL-2002.

XX

PF 18-JAN-2002; 2002WO-GB00228.

XX

PR 18-JAN-2001; 2001GB-0001312.

XX

PA (GLAX ) GLAXO GROUP LTD.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX

PI Blackstock WP, Hale RS, Prinjha R, Rowley A;

XX

DR WPI; 2002-599722/64.

DR N-PSDB; ABK90134.

XX  
PT Identifying modulators of Nogo or BACE activity for treating acute  
PT neuronal injuries, neoplastic or dysproliferative disorders, comprises  
PT providing and monitoring interaction between Nogo and BACE polypeptides  
PT -

XX  
PS Disclosure; Page 59-62; 68pp; English.

XX  
CC The present invention relates to a new method of identifying modulators  
CC of Nogo function or BACE activity. The method involves providing Nogo and  
CC BACE polypeptides capable of binding with each other, monitoring the  
CC interaction between these polypeptides, and determining if the test agent  
CC is a modulator of Nogo or BACE activity. The method is useful in treating  
CC acute neuronal injuries, such as spinal or head injury, stroke,  
CC peripheral nerve damage, and in neoplastic (e.g. glioblastomas,  
CC neuroblastomas), hyperproliferative or dysproliferative disorders (e.g.  
CC cirrhosis, psoriasis, keloid formation, fibrocystic conditions, tissue  
CC hypertrophy) of the central nervous system. The BACE polypeptide is  
CC useful in screening methods to identify agents that may act as modulators  
CC of BACE activity and in particular agents that may be useful in treating  
CC Nogo-associated diseases. The modulators of Nogo or BACE polypeptides,  
CC and the polynucleotide encoding the BACE polypeptide are useful in  
CC manufacturing a medicament for the treatment or prevention of disorders  
CC responsive to the modulation of Nogo activity, in alleviating the  
CC symptoms or improving the condition of a patient suffering from this  
CC disorder, in axon regeneration, or in preventing metastasis or spreading  
CC of a cancer. The polynucleotide may also be an essential component in  
CC assays, a probe, in recombinant protein synthesis, and in gene therapy  
CC techniques. The present amino acid sequence represents the human NogoA  
CC protein of the invention.

XX  
SQ Sequence 1192 AA;

Query Match 98.1%; Score 5810; DB 23; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 1e-293;  
Matches 1159; Conservative 4; Mismatches 14; Indels 16; Gaps 3;

Qy	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Db	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Qy	61	AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP	120
Qy	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Db	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Qy	181	SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP	226
		:   :	
Db	181	SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP	240
Qy	227	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	286
Db	241	LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	300



Qy	287	GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEVV	346
Db	301	GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNNQQELPTALTCLVKEDEVV	360
Qy	347	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	406
Db	361	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	420
Qy	407	DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF	466
Db	421	DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF	480
Qy	467	PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE	526
Db	481	PLLGDPSTENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE	540
Qy	527	EVVANMPEGLTPDLVQACESELNEVTGTIAIYETKMDLVQTSEVMQESLYPAAQLCPSF	586
Db	541	EVVANMPEGLTPDLVQACESELNEVTGTIAIYETKMDLVQTSEVMQESLYPAAQLCPSF	600
Qy	587	EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601	EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	705
Db	661	EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	720
Qy	706	MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	765
Db	721	MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	840
Qy	826	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLISSKTDSF SKLAREYTDLEVSHKSEI	885
Db	841	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLISSKTDSF SKLAREYTDLEVSHKSEI	900
Qy	886	ANAPDGAGSLPCTELPHDL SLKNIQPKVEEKISFSDDFS KNGSATS KVVLLPPDV SALGH	945
Db	901	ANAPDGAGSLPCTELPHDL SLKNIQPKVEEKISFSDDFS KNGSATS KVVLLPPDV SALA-	959
Qy	946	TQAEIESIVKPKVLEKEAEKKLPSDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV	1005
Db	960	TQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV	1019
Qy	1006	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1065
Db	1020	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1079
Qy	1066	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT	1125
Db	1080	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT	1139
Qy	1126	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178

Db 1140 |||L|L|A|L|I|S|L|F|S|V|P|V|I|Y|E|R|H|Q|A|I|D|H|Y|L|G|L|A|N|K|N|V|K|D|A|M|A|K|I|Q|A|K|I|P|G|L|K|R|K|A|E 1192

RESULT 8

AAU33228

ID AAU33228 standard; Protein; 1246 AA.

XX

AC AAU33228;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #3719.

XX

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US08656.

XX

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-611725/70.

XX

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -

XX

PS Claim 20; Page 737; 765pp; English.

XX

CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX

SQ Sequence 1246 AA;

Query Match 93.3%; Score 5526.5; DB 22; Length 1246;  
Best Local Similarity 93.0%; Pred. No. 5.9e-279;  
Matches 1123; Conservative 13; Mismatches 39; Indels 33; Gaps 10;

```
Qy      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db     42 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 101

Qy     61 AGLSAAVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSTVPAP 120
      |||
Db    102 AGLSAAVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSTVPAP 161

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    162 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 221

Qy    181 SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:      |:      : |||
Db    222 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 281

Qy    227 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    282 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 341

Qy    287 GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEV 346
      |||
Db    342 GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLVKEDEV 401

Qy    347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406
      |||
Db    402 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 461

Qy    407 DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466
      |||
Db    462 DKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 521

Qy    467 PLLDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 526
      |||
Db    522 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 581

Qy    527 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586
      |||
Db    582 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 641

Qy    587 EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESI KHEPENPPPPYE 646
      |||
Db    642 EESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESI KHEPENPPPPYE 701

Qy    647 EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 705
      |||
Db    702 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 761

Qy    706 MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMVLVKESLTETSFESMI 765
      |||
```

Db 762 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMI 821

Qy 766 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY 825  
 |||

Db 822 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY 881

Qy 826 SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLISSKTDSFSK LAREYTDLEVSHKSEI 885  
 |||

Db 882 SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLISSKTDSFSK LAREYTDLEVSHKSEI 941

Qy 886 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALGH 945  
 |||

Db 942 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALA- 1000

Qy 946 TQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV 1005  
 |||

Db 1001 TQAEIESIVKPKVLVKEAEKKLP SDTEKEDRSPSAIF SAELSKTSVVDLLYWRDIKKTGV 1060

Qy 1006 VFGAS-LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAY-- 1062  
 ||| : |||

Db 1061 VFGASAVFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAISG 1120

Qy 1063 -LESEVAISEELVQKYSNSALGHV-NCTIKELRR--LFLVDDLVD SLK-FAVLMWVFTY 1116  
 || : : | : ||| : ||| : | ||| : |||

Db 1121 NLESCLYLRELGSGRYSNSALGSMWNCTVKGNFRAPSF FSWMDLVD SLRSFAVLMWVFTY 1180

Qy 1117 VGALFNGLTLL-----ILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKI 1170  
 || | || : | || : |||

Db 1181 VGCL--GLMVLDTTGFWALNFISSSGSWLIYERHQAQIDHYLGLANKNVKDAMAKIQAKI 1238

Qy 1171 PGLKRKAE 1178  
 |||

Db 1239 PGLKRKAE 1246

# RESULT 9

ABU11573

ID ABU11573 standard; Protein; 983 AA.

XX

AC ABU11573;

XX

DT 12-FEB-2003 (first entry)

XX

DE Human MDDT polypeptide SEQ ID 520.

XX

KW MDDT; human; disease detection and treatment molecule polypeptide;  
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;  
 KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;  
 KW gene therapy; protein replacement therapy; cell proliferative disorder;  
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;  
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;  
 KW psoriasis; hepatitis.

XX

OS Homo sapiens.

XX

PN WO200279449-A2.

XX  
 PD 10-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US09944.  
 XX  
 PR 28-MAR-2001; 2001US-279619P.  
 PR 29-MAR-2001; 2001US-280067P.  
 PR 29-MAR-2001; 2001US-280068P.  
 PR 16-MAY-2001; 2001US-291280P.  
 PR 17-MAY-2001; 2001US-291829P.  
 PR 17-MAY-2001; 2001US-291849P.  
 PR 19-JUN-2001; 2001US-299428P.  
 PR 20-JUN-2001; 2001US-299776P.  
 PR 20-JUN-2001; 2001US-300001P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;  
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
 XX  
 DR WPI; 2003-058431/05.  
 DR N-PSDB; ABX34563.  
 XX  
 PT New purified disease detection and treatment molecule proteins and  
 PT polynucleotides, useful for diagnosing, treating or preventing cancers  
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis  
 PT or hepatitis -  
 XX  
 PS Claim 27; SEQ ID NO 520; 339pp + Sequence Listing; English.  
 XX  
 CC This invention describes a novel disease detection and treatment molecule  
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,  
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,  
 CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides  
 CC and the polypeptides of the invention can be used for gene therapy,  
 CC protein replacement therapy and are useful for treating a variety of  
 CC diseases or conditions. These polypeptides or polynucleotides are  
 CC particularly useful for diagnosing, treating or preventing cell  
 CC proliferative disorders (e.g. cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's  
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's  
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or  
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded  
 CC by ABU11450-ABU11845, described in the disclosure of the invention.  
 CC NOTE: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 983 AA;

Query Match 77.0%; Score 4560; DB 24; Length 983;  
 Best Local Similarity 97.1%; Pred. No. 7.8e-229;  
 Matches 919; Conservative 8; Mismatches 11; Indels 8; Gaps 3;

Qy	240	NLSTVL-----PTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVS	293
Db	39	NMNTLVICQYYPLKEHFKNVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVS	98
Qy	294	PKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVEDEVVSSEKAKD	353
Db	99	PKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVEDEVVSSEKAKD	158
Qy	354	SFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFAD	413
Db	159	SFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFAD	218
Qy	414	SLEQTNHEKDSSESSNDTSTPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPT	473
Db	219	SLEQTNHEKDSSESSNDTSTPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLGDP	278
Qy	474	SENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMP	533
Db	279	SENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMP	338
Qy	534	EGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATP	593
Db	339	EGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATP	398
Qy	594	SPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESI KHEPENPPPYEEAMSVSL	653
Db	399	SPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESI KHEPENPPPYEEAMSVSL	458
Qy	654	-KVSGIKKEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQP	712
Db	459	KKVSGIKKEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQP	518
Qy	713	VPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMI EYENKEK	772
Db	519	VPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMI EYENKEK	578
Qy	773	LSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIP LQMEELSTAVYSNDDLFI	832
Db	579	LSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIP LQMEELSTAVYSNDDLFI	638
Qy	833	SKEAQIRETETFSDDSPIEII DEFPTLISSKTSF SKLAREYTDLEVSHKSEIANAPDGA	892
Db	639	SKEAQIRETETFSDDSPIEII DEFPTLISSKTSF SKLAREYTDLEVSHKSEIANAPDGA	698
Qy	893	GSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALGH TQAEIES	952
Db	699	GSLPCTELPHDLSLKNIQPKVEEKISFSDDFS KNGSATS KVL LPPDV SALA-TQAEIES	757
Qy	953	IVKPKVLEKEAEKKLPSTDETEKEDRSPSAIF SADLGKTSVVDLLYWRDI KKTGVVFGASLF	1012
Db	758	IVKPKVLVKEAEKKLPSTDETEKEDRSPSAIF SAELSKTSVVDLLYWRDI KKTGVVFGASLF	817
Qy	1013	LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPPFRAYLESEVAISEE	1072
Db	818	LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPPFRAYLESEVAISEE	877
Qy	1073	LVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVFLMWVFTYVGALFNGLTLLI LALI	1132

Db 878 LVQKYSNSALGHVNCTIKELRRRLFVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALI 937  
 Qy 1133 SLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 Db 938 SLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 983

RESULT 10

AA95012

ID AAY95012 standard; Protein; 893 AA.

XX

AC AAY95012;

XX

DT 19-JUN-2000 (first entry)

XX

DE Human secreted protein vb22\_1, SEQ ID NO:64.

XX

KW Human; secreted protein; cancer; tumour; cardiovascular disorder;  
 KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;  
 KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;  
 KW neurodegenerative disease; asthma; contraceptive.

XX

OS Homo sapiens.

XX

PN WO200011015-A1.

XX

PD 02-MAR-2000.

XX

PF 24-AUG-1999; 99WO-US19351.

XX

PR 24-AUG-1998; 98US-0097638.

PR 24-AUG-1998; 98US-0097659.

PR 09-SEP-1998; 98US-0099618.

PR 28-SEP-1998; 98US-0102092.

PR 25-NOV-1998; 98US-0109978.

PR 23-DEC-1998; 98US-0113645.

PR 23-DEC-1998; 98US-0113646.

PR 23-AUG-1999; 99US-0379246.

XX

PA (ALPH-) ALPHAGENE INC.

XX

PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX

DR WPI; 2000-224657/19.

XX

PT New secreted or transmembrane proteins and polynucleotides encoding  
 PT them, useful for treating neurodegenerative disorders, autoimmune  
 PT diseases and cancer -

XX

PS Claim 73; Page 322-325; 357pp; English.

XX

CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),  
 CC and cDNA sequences encoding them (AAA23423-A23462). The secreted  
 CC proteins of the invention include those that are thought to be only  
 CC partially secreted, i.e., transmembrane proteins. The proteins of the  
 CC invention may exhibit one or more activities selected from the following:

CC cytokine activity; cell proliferation; differentiation; immune  
 CC modulation; haematopoiesis regulation; tissue growth activity;  
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic  
 CC and thrombolytic activity; anti-inflammatory activity; and tumour  
 CC inhibition activity. The proteins may be administered to patients as  
 CC vaccines, and the nucleotides may be used as part of a gene therapy  
 CC regime. Diseases or conditions that may be treated using the proteins or  
 CC nucleotides of the invention include autoimmune diseases; genetic  
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,  
 CC fungal and viral infections, especially HIV; multiple sclerosis;  
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;  
 CC insulin dependent diabetes mellitus; and allergic reactions such as  
 CC asthma and anaemia. They may also be used for treating wounds, burns,  
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic  
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may  
 CC additionally be useful as contraceptives. Nucleic acid sequences of the  
 CC invention may be used in chromosome mapping, and as a source of  
 CC diagnostic primers and probes. The present sequence represents one of the  
 CC 40 proteins of the invention.

XX

SQ Sequence 893 AA;

Query Match 74.3%; Score 4400; DB 21; Length 893;

Best Local Similarity 98.9%; Pred. No. 1.4e-220;

Matches 884; Conservative 1; Mismatches 7; Indels 2; Gaps 2;

QY	286	MGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEV	345
Db	1	MGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLVKEDEV	60
QY	346	VSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESK	405
Db	61	VSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESK	120
QY	406	VDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNI	465
Db	121	VDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNI	180
QY	466	FPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVT	525
Db	181	FPLLGDPSTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVT	240
QY	526	EEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPS	585
Db	241	EEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPS	300
QY	586	FESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPY	645
Db	301	FESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPY	360
QY	646	EEAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYS	704
Db	361	EEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYS	420
QY	705	EMAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMVLVKESLTETSFESM	764





FT		/note= "Acidic region"
FT	Region	31..57
FT		/note= "Region specifically described in claim 16"
FT	Modified-site	233
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	242..244
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	291
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	295
FT		/note= "Protein kinase C (PKC) site"
FT	Misc-difference	404
FT		/note= "Encoded by TTG"
FT	Modified-site	436
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	468..470
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	484
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	488
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	502
FT		/note= "Casein kinase II site"
FT	Modified-site	576
FT		/note= "Casein kinase II site"
FT	Modified-site	626
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	694..696
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	715
FT		/note= "Casein kinase II site"
FT	Modified-site	784
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	821
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	850
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	855
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	863
FT		/note= "Casein kinase II site"
FT	Modified-site	868
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	893
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	912..914
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	925..927
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	954
FT		/note= "PKC and casein kinase II sites"
FT	Modified-site	956
FT		/note= "PKC and casein kinase II sites"
FT	Domain	988..1023
FT		/label= Transmembrane_domain
FT		/note= "C-terminal hydrophobic region
FT		specifically described in claim 16"

FT Modified-site 1024  
 FT /note= "Protein kinase C (PKC) site"  
 FT Modified-site 1071..1073  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1073  
 FT /note= "Protein kinase C (PKC) site"  
 FT Modified-site 1089  
 FT /note= "Protein kinase C (PKC) site"  
 FT Domain 1090..1125  
 FT /label= Transmembrane domain  
 FT /note= "C-terminal hydrophobic region  
 specifically described in claim 16"  
 FT Modified-site 1141..1143  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1143  
 FT /note= "Protein kinase C (PKC) site"  
 FT Peptide 623..640  
 FT /note= "used as immunogen to generate antibody AS 472"  
 FT Peptide 762..1163  
 FT /note= "used as immunogen to generate antibody AS Bruna"  
 FT Inhibitory-site 542..722  
 FT Region 172..259  
 FT /note= "This region is not essential for inhibitory  
 activity"  
 FT Region 975..1162  
 FT /note= "This region is not essential for inhibitory  
 activity"  
 FT Region 976..1163  
 FT /note= "C-terminal common region found in Nogo A, B and  
 C isoforms"  
 XX  
 PN WO200031235-A2.  
 XX  
 PD 02-JUN-2000.  
 XX  
 PF 05-NOV-1999; 99WO-US26160.  
 XX  
 PR 06-NOV-1998; 98US-0107446.  
 XX  
 PA (SCHW/) SCHWAB M E.  
 PA (CHEN/) CHEN M S.  
 XX  
 PI Schwab ME, Chen MS;  
 XX  
 DR WPI; 2000-400052/34.  
 DR N-PSDB; AAD01173.  
 XX  
 PT Nogo proteins and nucleic acids useful for treating neoplastic  
 PT disorders of the central nervous system and inducing regeneration of  
 PT neurons -  
 XX  
 PS Claim 3; Fig 2A; 122pp; English.  
 XX  
 CC The present sequence is a rat Nogo A protein which is a  
 CC potent neural cell growth inhibitor and is free of all central nervous  
 CC system (CNS) myelin material with which it is natively associated. The  
 CC protein was derived from a cDNA generated by fusing R018U37-3, R1-3U21

CC cDNAs isolated from hexanucleotides-primed rat brain stem/spinal cord  
 CC library, and Olil8 cDNA from an oligo d(T)-primed rat oligodendrocyte  
 CC library. Nogo proteins and fragments displaying neurite growth  
 CC inhibitory activity are used in the treatment of neoplastic disease of  
 CC the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
 CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
 CC oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and  
 CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
 CC Therapeutics which promote Nogo activity can be used to treat or prevent  
 CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
 CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
 CC used to inhibit production of Nogo protein to induce regeneration of  
 CC neurons or to promote structural plasticity of the CNS in disorders where  
 CC neurite growth, regeneration or maintenance are deficient or desired.  
 CC The animal models can be used in diagnostic and screening methods for  
 CC predisposition to disorders and to screen for or test molecules which  
 CC can treat or prevent disorders or diseases of the CNS.  
 CC Note: The present sequence designated as SEQ ID NO: 2 is stated to  
 CC be the same as the sequence shown in Fig. 13 (see AAY71384) of the  
 CC specification. However, this sequence does not match the sequence given  
 CC in Fig. 13. SEQ ID numbers 35-42 are referred in claim 32 and  
 CC SEQ ID NO: 29 in disclosure of the specification. However, the  
 CC specification does not include sequences for these SEQ ID numbers.

XX

SQ Sequence 1163 AA;

Query Match 72.5%; Score 4296.5; DB 21; Length 1163;  
 Best Local Similarity 74.0%; Pred. No. 4.8e-215;  
 Matches 885; Conservative 104; Mismatches 156; Indels 51; Gaps 19;

Qy	1	MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDDEE-EEEEEEEEDEDEDLEELEVLERK	58
Db	1	MEDIDQSSLVSSSTDSPRPAPPAFKYQFVTEPEDEDEEEEEDEDEDLEELEVLERK	60
Qy	59	PAAGLSAAPVPTAPAAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP	118
Db	61	PAAGLSAAAVP--PAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Qy	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Db	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Qy	179	RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSL	224
Db	167	RGSGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL	226
Qy	225	SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRLTEFSELEYS	284
Db	227	SPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPVNRDLAEFSELEYS	286
Qy	285	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDE	344
Db	287	EMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KEDR	340
Qy	345	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE	403
Db	341	VVSPEKTMDFNEMQMSVAVPVREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANVE	396



XX  
 AC ABB81074;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Rat neurotransmitter receptor protein Nogo-A.  
 XX  
 KW Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;  
 KW central nervous system; peripheral nervous system; tranquillizer; Nogo;  
 KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;  
 KW nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;  
 KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;  
 KW neurotransmitter receptor; rat; receptor.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN US2002072493-A1.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 28-JUN-2001; 2001US-0893348.  
 XX  
 PR 19-MAY-1998; 98IL-0124500.  
 PR 21-JUL-1998; 98WO-US14715.  
 PR 22-DEC-1998; 98US-0218277.  
 PR 19-MAY-1999; 99US-0314161.  
 XX  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 PI Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonego A;  
 PI Moalem G;  
 XX  
 DR WPI; 2002-607255/65.  
 DR N-PSDB; ABN86600.  
 XX  
 PT Promoting nerve regeneration and preventing neuronal degeneration in  
 PT the central/peripheral nervous system from injury/disease, comprises  
 PT administering nervous system-specific activated T cells/antigen, or  
 PT analogs/peptides -  
 XX  
 PS Example 5; Page 44-47; 93pp; English.  
 XX  
 CC The invention relates to promoting nerve regeneration or conferring  
 CC neuroprotection and preventing or inhibiting neuronal degeneration in the  
 CC central/peripheral nervous system (NS). The method involves administering  
 CC NS-specific activated T cells, NS-specific antigen, its analogue or its  
 CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or  
 CC combinations. The method is useful for promoting nerve regeneration and  
 CC preventing neuronal degeneration in central/peripheral nervous system  
 CC from injury/disease, where the injury is spinal cord injury, blunt  
 CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or  
 CC damages caused by surgery such as tumour excision. The disease is not an  
 CC autoimmune disease or neoplasm. The disease results in a degenerative  
 CC process occurring in either gray or white matter or both. The disease  
 CC is diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's  
 CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,  
 CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and

CC vitamin deficiency, intervertebral disc herniation, prion diseases such  
 CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral  
 CC neuropathies associated with various diseases, including but not limited  
 CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute  
 CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary  
 CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption  
 CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-  
 CC pathies, complications of various drugs (e.g., metronidazole) and toxins  
 CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia  
 CC telangectasia, Friedreich's ataxia, amyloid polyneuropathies,  
 CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's  
 CC disease, or lipoproteinemia. The present sequence represents the rat  
 CC neurotransmitter receptor protein Nogo-A, an example of NS-specific  
 CC antigen.

XX

SQ Sequence 1163 AA;

Query Match 72.5%; Score 4296.5; DB 23; Length 1163;  
 Best Local Similarity 74.0%; Pred. No. 4.8e-215;  
 Matches 885; Conservative 104; Mismatches 156; Indels 51; Gaps 19;

Qy	1	MEDLDQSPLVSSS-DSPPRQPAPKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK	58
		:	
Db	1	MEDIDQSSLVSSSTDSPPRPPAPKYQFVTEPEDEEDEEEEEDEEDEDLEELEVLERK	60
Qy	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP	118
Db	61	PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPA---P	115
Qy	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Db	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Qy	179	RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSSL	224
		:	
Db	167	RGSGSVDETFLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL	226
Qy	225	SPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS	284
		:       :   :         :   :	
Db	227	SPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYS	286
Qy	285	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDE	344
		:     :     :     :     :	
Db	287	EMGSSFKGSPKGESAILVENTKEEIVRSKDKED-LVCSAALHSPQESPVG-----KEDR	340
Qy	345	VVSSEKAKDSFNEKRVAEAPMREEYADFKPFERVVEVKDSKEDS-DMLAAGGKIESNLE	403
		:     :     :     :     :	
Db	341	VVSPEKTMDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANVE	396
Qy	404	SKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIAT	463
		:           :           :           :           :	
Db	397	SKVDRKCLEDSEKSLGKDSSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA	455
Qy	464	NIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTK	523
		:     :	
Db	456	NTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLSK	514





KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
KW structural plasticity; screening; mutant; mutein.

XX

OS Rattus sp.

XX

PN WO200031235-A2.

XX

PD 02-JUN-2000.

XX

PF 05-NOV-1999; 99WO-US26160.

XX

PR 06-NOV-1998; 98US-0107446.

XX

PA (SCHW/) SCHWAB M E.

PA (CHEN/) CHEN M S.

XX

PI Schwab ME, Chen MS;

XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic  
PT disorders of the central nervous system and inducing regeneration of  
PT neurons -

XX

PS Example; Page -; 122pp; English.

XX

CC The patent relates to neurite growth inhibitor Nogo which is free of  
CC all central nervous system (CNS) myelin material with which it is  
CC natively associated. Nogo proteins and fragments displaying neurite  
CC growth inhibitory activity are used in the treatment of neoplastic  
CC disease of the CNS e.g. glioma, glioblastoma, medulloblastoma,  
CC craniopharyngioma, ependyoma, pinealoma, haemangioblastoma, acoustic  
CC neuroma, oligodendroglioma, menagioma, neuroblastoma or retinoblastoma  
CC and degenerative nerve diseases e.g. Alzheimer's and Parkinson's  
CC diseases. Therapeutics which promote Nogo activity can be used to treat  
CC or prevent hyperproliferative or benign dysproliferative disorders e.g.  
CC psoriasis and tissue hypertrophy. Ribozymes or antisense Nogo nucleic  
CC acids can be used to inhibit production of Nogo protein to induce  
CC regeneration of neurons or to promote structural plasticity of the CNS  
CC in disorders where neurite growth, regeneration or maintenance are  
CC deficient or desired. The animal models can be used in diagnostic and  
CC screening methods for predisposition to disorders and to screen for or  
CC test molecules which can treat or prevent disorders or diseases of the  
CC CNS. The present sequence is a truncated form of rat Nogo A protein shown  
CC in AAY71310, which is used in the construction of mutant Nogo-A. Nogo-A  
CC is composed of His-tag/T7-tag/vector/Nogo-A sequence aa 1-1162.  
CC Nogo A deletion mutants were used for mapping the inhibitory sites of  
CC Nogo protein. Major inhibitory region was identified in the  
CC Nogo A sequence from amino acids 172-974, particularly amino acids  
CC 542-722. In addition, N-terminal region 1-171 was found to be inhibitory  
CC to NIH 3T3 fibroblast spreading.  
CC Note: The present sequence is not given in the specification but is  
CC derived from rat Nogo A sequence shown in AAY71310. SEQ ID numbers 35-42  
CC are referred in claim 32 and SEQ ID NO: 29 in disclosure of the

CC specification. However, the specification does not include sequences for  
CC these SEQ ID numbers.

XX

SQ Sequence 1162 AA;

Query Match 72.5%; Score 4294.5; DB 21; Length 1162;  
Best Local Similarity 74.1%; Pred. No. 6.1e-215;  
Matches 885; Conservative 103; Mismatches 156; Indels 51; Gaps 19;

```
Qy      1 MEDLDQSPLVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58
      |||:|||| |||| |||| |||| |||| |||| ||||:|:|:|||||||
Db      1 MEDIDQSSLVSSSTDSPPRPPAPFKYQFVTEPEDEEDEEEEEDEDEDLEEEVLERK 60

Qy     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSSTVP 118
      ||||| || || |||:|:| ||||| ||||| |||:|:|
Db     61 PAAGLSAAAVP--PAAAAPLLDfSSDSVFPAPRGPLPAAPPAAPERQPSWERSPA--P 115

Qy    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178
      ||| ||| ||||| ||||| ||||| ||| || ||||| |||||
Db    116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166

Qy    179 RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSL 224
      ||| : |||| |||||:|:||||||| |||
Db    167 RGSGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL 226

Qy    225 SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 284
      ||| |||| ||||| | :|:|:| :|:||||: |:| :|:| |||||
Db    227 SPLSTVSFKEHGYLGNL SAVSSSEG TIEETLNEASKELPERATNPFVNRDLAEFSELEYS 286

Qy    285 EMGSSFSVSPKAESAVIVANPREEI IVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDE 344
      ||||| ||| |||:| | :|:|:|:|:|:| | : || || |||
Db    287 EMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KEDR 340

Qy    345 VVSSEKAKDSFNEKRAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE 403
      ||| || | ||| :|:| | |:|||||||:|:| |||:| | |:| | :|:|
Db    341 VVSPEKTMDFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANVE 396

Qy    404 SKVDKKCFADSLEQTNHEKDSSESSNDTSFSPSTPEGIKDRSGAYITCAPFNPAATESIAT 463
      |||:| ||||| : |||| |:| ||||| :| | ||||| | :|||
Db    397 SKVDRKCIEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA 455

Qy    464 NIFPILLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTK 523
      | ||||| |||| |||||:|:|:| || ||||| || ||||| |||:|
Db    456 NTFPILLEDHTSENKTDEKKIEERKAQII TEK-TSPKTSNPFLVAVQDSEADYVTTDTLSK 514

Qy    524 VTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLC 583
      ||| |:| ||||| ||||| ||||| |||||:|:| ||||| :| ||||| |||
Db    515 VTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLC 574

Qy    584 PSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPENP 642
      ||||:| ||||| ||||| ||||| :|:| ||||| ||||| |:|:| |||||
Db    575 PSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVPLEAPPPVSYSIKLEPENP 634

Qy    643 PPYEEAMSVSLKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSD 702
      |||||:|:| | || ||||: ||:| ||||| ||||| |||:|:|:|
Db    635 PPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSDFSN 694
```

Qy 703 YSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFE 762  
 |||:| | : ||:|:|||||||:|||||||:| | | :| |||:||||| | |  
 Db 695 YSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSI PEVPQTQEEAVMLMKESLTEVS-E 753  
 Qy 763 SMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELST 822  
 :: :: :|:| | | | |||||: :| :||| ::: ||:|||| | ||| :|  
 Db 754 TVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEFNT 811  
 Qy 823 AVYSNDDLFISKEAQIRETETFSDDSPIEI IDEFPTLISSKTDSFSKLAREYTDLEVSHK 882  
 |:||||| ||| :|:|:|||||||:|:| | | ||:||||||| |  
 Db 812 AIYSNDDLLSSKEDKIKESETFSDDSPIEI IDEFPTFVSAKDDS-PKLAKEYTDLEVSDK 870  
 Qy 883 SEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVS 942  
 ||||| | ||||| ||| ||| ||| ::: ||:|:| | : | | : |||  
 Db 871 SEIANIQSGADSLPCLELPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNVSA 928  
 Qy 943 LGHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKK 1002  
 | | | : |||| | | |||||:|:| | | |||||:|:| |  
 Db 929 L-EPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKK 987  
 Qy 1003 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY 1062  
 |||||:|:| | | |||||:|:| | | |||||:|:| | | |||||:|:| |  
 Db 988 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY 1047  
 Qy 1063 LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFN 1122  
 |||||:|:| | | |||||:|:| | | |||||:|:| | | |||||:|:| |  
 Db 1048 LESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFN 1107  
 Qy 1123 GLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKA 1177  
 |||||:|:| | | |||||:|:| | | |||||:|:| | | |||||:|:| |  
 Db 1108 GLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKA 1162

# RESULT 14

AA71384

ID AA71384 standard; Protein; 1163 AA.

XX

AC AA71384;

XX

DT 02-NOV-2000 (first entry)

XX

DE Alternative version of rat neurite growth inhibitor Nogo A.

XX

KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS;  
 KW central nervous system; neoplastic disease; antiproliferative; glioma;  
 KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
 KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
 KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
 KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
 KW structural plasticity; screening.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT Inhibitory-site 1..171

FT /note= "Inhibits NIH 3T3 fibroblast spreading"

FT Modified-site 30

FT		/note= "Casein kinase II site"
FT	Region	31..58
FT		/note= "Acidic region"
FT	Modified-site	233
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	242..244
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	291
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	295
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	436
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	468..470
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	484
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	488
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	502
FT		/note= "Casein kinase II site"
FT	Modified-site	576
FT		/note= "Casein kinase II site"
FT	Modified-site	626
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	694..696
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	715
FT		/note= "Casein kinase II site"
FT	Modified-site	784
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	821
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	850
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	855
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	863
FT		/note= "Casein kinase II site"
FT	Modified-site	868
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	893
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	912..914
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	925..927
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	954
FT		/note= "PKC and casein kinase II sites"
FT	Modified-site	956
FT		/note= "PKC and casein kinase II sites"
FT	Domain	988..1023
FT		/label= Transmembrane_domain
FT		/note= "C-terminal hydrophobic region"
FT	Modified-site	1024
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	1071..1073

FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1073  
 FT /note= "Protein kinase C (PKC) site"  
 FT Modified-site 1089  
 FT /note= "Protein kinase C (PKC) site"  
 FT Domain 1090..1125  
 FT /label= Transmembrane\_domain  
 FT /note= "C-terminal hydrophobic region"  
 FT Modified-site 1141..1143  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1143  
 FT /note= "Protein kinase C (PKC) site"  
 FT Peptide 623..640  
 FT /note= "used as immunogen to generate antibody AS 472"  
 FT Peptide 762..1163  
 FT /note= "used as immunogen to generate antibody AS Bruna"  
 FT Inhibitory-site 542..722  
 FT Region 172..259  
 FT /note= "This region is not essential for inhibitory  
 FT activity"  
 FT Region 975..1162  
 FT /note= "This region is not essential for inhibitory  
 FT activity"  
 FT Region 976..1163  
 FT /note= "C-terminal common region found in Nogo A, B and  
 FT C isoforms"  
 FT Misc-difference 223  
 FT /label= Unknown  
 FT /note= "There is Leu at this position in the  
 FT sequence shown in AAY71310"  
 FT Misc-difference 404  
 FT /note= "There is Ile at this position in the  
 FT sequence shown in AAY71310"  
 FT Misc-difference 469  
 FT /label= Unknown  
 FT /note= "There is Lys at this position in the  
 FT sequence shown in AAY71310"  
 FT Misc-difference 661  
 FT /note= "There is Asn at this position in the  
 FT sequence shown in AAY71310"  
 FT Misc-difference 820  
 FT /note= "There is Leu at this position in the  
 FT sequence shown in AAY71310"  
 XX  
 PN WO200031235-A2.  
 XX  
 PD 02-JUN-2000.  
 XX  
 PF 05-NOV-1999; 99WO-US26160.  
 XX  
 PR 06-NOV-1998; 98US-0107446.  
 XX  
 PA (SCHW/) SCHWAB M E.  
 PA (CHEN/) CHEN M S.  
 XX  
 PI Schwab ME, Chen MS;  
 XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic  
PT disorders of the central nervous system and inducing regeneration of  
PT neurons -

XX

PS Claim 3; Fig 13; 122pp; English.

XX

CC The present sequence is an alternative version of rat Nogo A protein  
CC which is a potent neural cell growth inhibitor and is free of all  
CC central nervous system (CNS) myelin material with which it is  
CC natively associated. Nogo proteins and fragments displaying  
CC neurite growth inhibitory activity are used in the  
CC treatment of neoplastic disease of the CNS  
CC e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
CC oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and  
CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
CC Therapeutics which promote Nogo activity can be used to treat or prevent  
CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
CC used to inhibit production of Nogo protein to induce regeneration of  
CC neurons or to promote structural plasticity of the CNS in disorders where  
CC neurite growth, regeneration or maintenance are deficient or desired.  
CC The animal models can be used in diagnostic and screening methods for  
CC predisposition to disorders and to screen for or test molecules which  
CC can treat or prevent disorders or diseases of the CNS.  
CC Note: The present sequence is an alternative version of the  
CC Nogo A sequence shown in Fig. 2A (see AAY71310).  
CC SEQ ID numbers 35-42 are referred in claim 32 and SEQ ID NO: 29  
CC in disclosure of the specification. However the specification does not  
CC include sequences for these SEQ ID numbers.

XX

SQ Sequence 1163 AA;

Query Match 72.4%; Score 4286.5; DB 21; Length 1163;  
Best Local Similarity 74.1%; Pred. No. 1.6e-214;  
Matches 886; Conservative 104; Mismatches 155; Indels 51; Gaps 19;

```
QY      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58
      |||: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEDEDLEELEVLERK 60

QY     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118
      ||||| ||| ||| |||: ||| : ||||| ||||| ||||| ||||| : || : |
Db     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115

QY    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178
      ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166

QY    179 RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSL 224
      ||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    167 RGSGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASXPSL 226

QY    225 SPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 284
      |||| ||||| ||||| | :|||:| :|||:| :| :||| ||||| |||||
```

Db	227	SPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYS	286
Qy	285	EMGSSFSVSPKAESAIVANPREEIIVKNKDEEEKLVSNINILHXQQELPTALTCLKVKEDE	344
Db	287	EMGSSFKGSPKGESAILVENTKKEEIVRSKDKED-LVCSAALHSPQESPVG-----KEDR	340
Qy	345	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE	403
Db	341	VVSPEKTMDFNEMQMSVVAVPREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANVE	396
Qy	404	SKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIAT	463
Db	397	SKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA	455
Qy	464	NIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTK	523
Db	456	NTFPLLEDHTSENXTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLSK	514
Qy	524	VTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLC	583
Db	515	VTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLC	574
Qy	584	PSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSSPLEA-SSVNYESIKHEPENP	642
Db	575	PSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSYSIKLEPENP	634
Qy	643	PPYEEAMSVSLKVSGIKKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFS	702
Db	635	PPYEEAMNVALKALGTKEGIKEPESFMAAVQETEAPYISIIACDLIKETKLSTEPSPDFSN	694
Qy	703	YSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFE	762
Db	695	YSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVS-E	753
Qy	763	SMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELST	822
Db	754	TVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLOMEEFNT	811
Qy	823	AVYSNDDLFIKSEAQIRETETFSDDSPIEIIDFPTLISSKTDSFSKLAREYTDLEVSHK	882
Db	812	AIYSNDDLSSSKEDKIKESETFSDDSPIEIIDFPTFVSAKDDS-PKLAKEYTDLEVSDK	870
Qy	883	SEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVS	942
Db	871	SEIANIQSGADSLPCLLEPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNVSA	928
Qy	943	LGHTQAEIESIVKPKVLEKEAEKKLPSPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKK	1002
Db	929	L-EPQTEMGSI VKSKSLTKEAEKKLPSPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKK	987
Qy	1003	TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1062
Db	988	TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1047
Qy	1063	LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFN	1122
Db	1048	LESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFN	1107







Db 397 SKVDRKCIEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDDSSRAYITCASFT-SATESTTA 455  
 Qy 464 NIFPILLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTK 523  
 | | | | | | | | | | | | | | | | : | | | | : | | | | | | | | | | | | : |  
 Db 456 NTFPILLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFVLVAVQDSEADYVTTDTLSK 514  
 Qy 524 VTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLC 583  
 | | | | : | | | | | | | | | | | | | | | | | : | | | | | : | | | | | | |  
 Db 515 VTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLC 574  
 Qy 584 PSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPENP 642  
 | | | | : | | | | | | | | | | | | | | | | : | | | | : | | | | | | | | | |  
 Db 575 PSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENP 634  
 Qy 643 PPYEEAMSVSLKVSIGKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSD 702  
 | | | | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | : | | | :  
 Db 635 PPYEEAMNVALKALGTKEGIEKESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSN 694  
 Qy 703 YSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFE 762  
 | | | : | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 695 YSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSI PEVPQTQEEAVMLMKESLTEVS-E 753  
 Qy 763 SMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELST 822  
 : : : : : | : | | | | | | | | | | : | : | | | : : : | | : | | | | | | | : |  
 Db 754 TVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEFNT 811  
 Qy 823 AVYSNDDLFIKESQIRETETFSDDSPIEIIDFPPTLISSKTDSFSKLAREYTDLEVSHK 882  
 | : | | | | | | | | : | : | : | | | | | | | | | | | | | | | | | | | | | |  
 Db 812 AIYSNDDLSSKEDKIKESETFSDDSPIEIIDFPPTFVSAKDDS-PKLAKEYTDLEVSDK 870  
 Qy 883 SEIANAPDGAGSLPCTELPHDLNLSKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDPVSA 942  
 | | | | | | | | | | | | | | | | | | | | : : : | | : | | : | | : | : | | |  
 Db 871 SEIANIQSGADSLPCLELPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNVSA 928  
 Qy 943 LGHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSA DLGKT 989  
 | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 929 L-EPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKT 974

Search completed: January 22, 2004, 16:36:36  
 Job time : 84.555 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:31:15 ; Search time 28.5792 Seconds  
 (without alignments)  
 1744.001 Million cell updates/sec

Title: US-09-830-972-29  
 Perfect score: 5923  
 Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	931	15.7	199	2	US-08-700-607-1
2	787.5	13.3	776	2	US-08-700-607-5
3	705	11.9	356	2	US-08-700-607-6
4	688	11.6	208	2	US-08-700-607-7
5	671	11.3	267	2	US-08-700-607-8
6	541.5	9.1	168	4	US-09-149-476-563
7	513	8.7	241	2	US-08-700-607-3
8	316	5.3	8991	4	US-08-714-741-32
9	285	4.8	92	4	US-09-149-476-411
10	276.5	4.7	1786	3	US-08-973-462-8
11	267.5	4.5	1601	4	US-09-345-473E-40

12	265	4.5	1596	4	US-08-978-277A-4	Sequence 4, Appli
13	263.5	4.4	2409	6	5180808-2	Patent No. 5180808
14	256	4.3	1780	1	US-08-769-309A-5	Sequence 5, Appli
15	256	4.3	1780	3	US-08-994-570-5	Sequence 5, Appli
16	238	4.0	1346	2	US-08-635-121-2	Sequence 2, Appli
17	238	4.0	1346	4	US-08-978-277A-2	Sequence 2, Appli
18	237.5	4.0	688	3	US-09-141-047-8	Sequence 8, Appli
19	237.5	4.0	1270	4	US-07-757-022B-44	Sequence 44, Appl
20	237.5	4.0	1311	4	US-07-757-022B-42	Sequence 42, Appl
21	237.5	4.0	1313	4	US-07-757-022B-142	Sequence 142, App
22	237.5		1314	4	US-07-757-022B-50	Sequence 50, Appl
23	237.5	4.0	1320	4	US-07-757-022B-46	Sequence 46, Appl
24	237.5	4.0	1320	4	US-07-757-022B-60	Sequence 60, Appl
25	237.5	4.0	1354	4	US-07-757-022B-48	Sequence 48, Appl
26	237.5	4.0	1361	4	US-07-757-022B-40	Sequence 40, Appl
27	237.5	4.0	1363	4	US-07-757-022B-52	Sequence 52, Appl
28	237.5	4.0	1404	4	US-07-757-022B-2	Sequence 2, Appli
29	237.5	4.0	1404	4	US-07-757-022B-62	Sequence 62, Appl
30	234.5	4.0	941	4	US-07-757-022B-14	Sequence 14, Appl
31	234.5	4.0	1022	4	US-07-757-022B-84	Sequence 84, Appl
32	234.5	4.0	1038	4	US-07-757-022B-74	Sequence 74, Appl
33	234.5	4.0	1049	4	US-07-757-022B-58	Sequence 58, Appl
34	234.5	4.0	1140	4	US-07-757-022B-104	Sequence 104, App
35	233.5	3.9	2137	4	US-09-134-001C-4463	Sequence 4463, Ap
36	226.5	3.8	630	3	US-08-973-462-9	Sequence 9, Appli
37	225	3.8	783	6	5231168-2	Patent No. 5231168
38	221	3.7	1087	4	US-09-914-259-12	Sequence 12, Appl
39	221	3.7	2482	4	US-09-252-991A-16967	Sequence 16967, A
40	220	3.7	1481	2	US-08-616-844-40	Sequence 40, Appl
41	220	3.7	1481	2	US-08-599-654-40	Sequence 40, Appl
42	220	3.7	1481	3	US-08-944-868A-40	Sequence 40, Appl
43	220	3.7	1400X7		-08-944-423A-40	Sequence 40, Appl
44	220	3.7	1481	3	US-08-944-496-40	Sequence 40, Appl
45	216.5	3.7	1848	3	US-08-296-791-6	Sequence 6, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-700-607-1

; Sequence 1, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304



```

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307307
US-08-700-607-5

```

```

Query Match          13.3%; Score 787.5; DB 2; Length 776;
Best Local Similarity 31.9%; Pred. No. 3.1e-40;
Matches 229; Conservative 91; Mismatches 190; Indels 207; Gaps 25;

```

```

Qy      588 ESEATPSPVLPDI--VMEAPLNSA-----VPSAGASVIQPSSSPLE---ASSVNY---- 632
      |  || | |||: : | : | : | : | : | : |
Db      141 EELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKKILADPLDQMKAAYKYIDIT 200

Qy      633 --ESIKHEPENPPPYEEA-----MSVSLKVSGIKEEIKEPENINAAL-----QET 675
      | :||: :: | | : :|| | :| | :| | : : : |
Db      201 RPEEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDK-PAPVEGKI IKDHLLEEST 259

Qy      676 EAPYISIACDL-----IKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELVED 722
      |||| | | : | | | : : : : | : :
Db      260 FAPYID---DLSEEQRRAPQITTPVKITLTEIEPSVE-----TTTQEKTPKQDICLK 309

Qy      723 SSPDSEPVDLFS---DDS-----IPDVPQKQ-----DETVMLVKESLTETSF 761
      |||: | | || | : | | : : || :
Db      310 PSPDTVPTVTVSEPEDDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEA----- 363

Qy      762 ESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELS 821

```

```

      : || | : : |
Db      364 -KGLSYETAENPRPVGQLADRP-----EVKARSGPPTIPSPLDHEA 403
Qy      822 TAVYSNDDLFI SKEAQIRETETFS DSSPIEI IDEFP---TLISSKTDSFS----- 868
      :: | | | ||:: | | : | ||
Db      404 SSAESGD-----SEIELVSEDPMAAEDALPSGYVSFGHVGGPPPPSP 444
Qy      869 -----KLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSD 921
      : || : | : : | : : | : | | | | |
Db      445 ASPSIQYSILREEREAE L DSELI IESCDASSAS-----EESPKREQDSPMKPSALD 496
Qy      922 DF-----SKNGSATSKVLL-----LPPDV SALGHTQAEIESIVKP 956
      | : | | | | | | | | | | | | | | | | | | | | |
Db      497 AIREETGVRAERAPSRRLAEPGSFLDYPSTEPQGPPELPPGDGAL-----EPETPMLP 551
Qy      957 KVLEKEAEKKLPSDTEKEDRSPSA-----IFSADLGKTSVVDLLYWRDIK 1001
      : | : : ||: | : | | | | | | | | | | | | | | |
Db      552 -----RKPEEDSSSNQSPAATKGPGPLGPGAPPPLLF---LNKQKAIDLLYWRDIK 599
Qy      1002 KTG VVFGASL FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1061
      : ||: ||: | | | | | | | | | | | | | | | | | | | | | | | |
Db      600 QTGIVFGSFL LLLFSLTQFSVVS VVAYLALALALSATISFRIYKSVLQAVQKTDEGHPFKA 659
Qy      1062 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1121
      || | | : : | : ||: : : | | | | | | | | | | | | | | | | |
Db      660 YLELEITLSQEIQKYTDCLQFYVNSTLKE LRRLFLVQDLVDLSLKFAVLMWLLTYVGALF 719
Qy      1122 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178
      || | | | | : : | : ||: ||: | | | | | | | | | | | | | | | |
Db      720 NGLTLLLMAVVSMTLPVVYVKHQAQIDQYLG LVRTHINAVVAKIQAKIPGAKRHAE 776

```

RESULT 3

US-08-700-607-6

; Sequence 6, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/700,607

```

; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307309
US-08-700-607-6

```

```

Query Match          11.9%; Score 705; DB 2; Length 356;
Best Local Similarity 42.6%; Pred. No. 1.2e-35;
Matches 168; Conservative 55; Mismatches 86; Indels 85; Gaps 11;

```

```

Qy      803 EVSTLSKKEKIPLQMEELSTAVYSNDDLFIKSEAQIRETETFS DSSPIEIIDEFPTLISS 862
      : | | : : | | : : | | : | : | | | | : : | : :
Db      30 QYSILREEREAE L DSELI ---IESCDASSASEESPKRE---QDSPPMK-----PSALDA 77

Qy      863 KTDFSFKLAREYTDLEVSHKSEIANAPDGAGSL---PCTELPHDLSLKNIQPKVEEKISF 919
      | | | : : : : | | | | | | | | | | | | | | | | |
Db      78 -----IREETGVRAEERAPSRRGLAEFGSFLDYPSTE-----PQGPGE----- 115

Qy      920 SDDFSKNGSATS K V L L L P P D V S A L G H T Q A E I E S I V K P K V L E K E A E K K L P S D T E K E D R S P S 979
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      116 -----LPPGDGAL-----EPETPMLP-----RKPEEDSSSNQSPA 145

Qy      980 A-----IFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLT VFSIVS 1024
      | : | | | : | | | : | | | | | | | | | | | | | |
Db      146 ATKGPGLPGGAPPPLLF---LNKQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVS 202

Qy      1025 VTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGH 1084
      | | | : | | | | | | | | | | | | | | | | | | | |
Db      203 VVAYLALAALSATISFRIYKSVLQAVQKTDEGHPPFKAYLELEITLSQEIQKYTDCLQFY 262

Qy      1085 VNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERH 1144
      | | | : | | | | | | | | | | | | | | | | | | | |
Db      263 VNSTLKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMTLPVVVYVKH 322

Qy      1145 QAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178
      | | | | | | | : : : | | | | | | | | | |
Db      323 QAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHAE 356

```

```

RESULT 4
US-08-700-607-7
; Sequence 7, Application US/08700607
; Patent No. 5858708

```



```

; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307311
US-08-700-607-7

```

```

Query Match          11.6%; Score 688; DB 2; Length 208;
Best Local Similarity 67.5%; Pred. No. 5.9e-35;
Matches 129; Conservative 32; Mismatches 30; Indels 0; Gaps 0;

```

```

Qy      988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV I 1047
      | : : ||||| : || : || : | || ||| || : || | | ||||| | :
Db      18 KSQAIDLLYWRDIKQTGIVFGSFLLLLSLTQFSVSVSVVAYLALAALSATISFRIYKSVL 77

Qy     1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKF 1107
      || : || : ||||| : |||| | : : | : ||| : : : || | : ||||| || |||||
Db      78 QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKF 137

Qy     1108 AVLMMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167
      ||||| : ||||| ||||| : ||| : || : || : || : ||||| |||| : : : ||||
Db     138 AVLMMWLLTYVGALFNGLTLLLMVVMFTLPVVVYVKHQAQIDQYLGVLVRTHINAVVAKIQ 197

```

Qy 1168 AKIPGLKRKAE 1178  
      ||||| || ||  
Db 198 AKIPGAKRHAE 208

RESULT 5

US-08-700-607-8

; Sequence 8, Application US/08700607  
; Patent No. 5858708  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Hillman, Jennifer L.  
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: U.S.  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/700,607  
; FILING DATE: Filed Herewith  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0114 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 281046

US-08-700-607-8

Query Match 11.3%; Score 671; DB 2; Length 267;  
Best Local Similarity 66.3%; Pred. No. 9.4e-34;  
Matches 124; Conservative 34; Mismatches 29; Indels 0; Gaps 0;

Qy 988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 1047  
      |: :|||||||:|:|:|: | | | | |:| | | | | | | | | | |:  
Db 9 KSQAIDLLYWRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFRIYKSVL 68

Qy	1048	QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDLDLVDLSKF	1107
		:     :         :           :   :   :   :     :     :	
Db	69	QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQLYVNSTLKELRRLFLVQDLVDLSKF	128
Qy	1108	AVLMWVFITYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ	1167
		:                     :   :   :   :   :     :         :         :	
Db	129	AVLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVVYVKHQAVDQYLGLVRTHINTVVAKIQ	188
Qy	1168	AKIPGLK	1174
		:	
Db	189	AKIPGAR	195

## RESULT 6

US-09-149-476-563

; Sequence 563, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: PZ002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,503

; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594

```
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
```

```
Query Match          9.1%; Score 541.5; DB 4; Length 168;
Best Local Similarity 60.5%; Pred. No. 4.2e-26;
Matches 101; Conservative 34; Mismatches 31; Indels 1; Gaps 1;
```

```
Qy      1013 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEE 1072
          :|||  ||::|| :|: ||||| ||||| ||||: |||: ||||: |||: :| |
Db      1 MLLSLAAPSVISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSE 60

Qy      1073 LVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALI 1132
          | |:|: |:| :| : ||||: ||||| || ||: ||||: |||: ||||| :
Db      61 AFHNYMNAAMVHINRALKLIIRLFLVEDLVDLSKLAVFMWLMITYVGAVFNGITLLILAE 120
```



Db	20	AEPSAPGGGGSPGACPALGTKSCSSSCAVHDLIXWRDVKKTGfVFGTTLIMLLSLAAFSV	79
Qy	1023	VSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSAL	1082
Db	80	ISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAM	139
Qy	1083	GHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGFALFNGLTLLILALISLFSVPVIYE	1142
Db	140	VHINRALKLIIRLFLVEDLVDLSLKLAVFMWLMTYVGAVFNGITLLILAELLIXSVPIVYX	199
Qy	1143	RHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRK	1176
Db	200	KY-----KVPSKTPWNROK	213

## RESULT 8

US-08-714-741-32

; Sequence 32, Application US/08714741

; Patent No. 6500613

## ; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Crain, Marilyn J.

; APPLICANT: Hollingshead, Susan

; APPLICANT: Tart, Rebecca

; APPLICANT: Brooks-Walter, Alexis

10 TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,

; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES.

; TITLE OF INVENTION: PORTIONS AND PRODUCTS

```
; NUMBER OF SEQUENCES: 47
```

## ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford, P.C.

STREET: 530 Fifth Avenue

; CITY: New York

STATE: New York

COUNTRY: U.S.

ZIP: 10036

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

;      COMPUTER:  IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

;      SOFTWARE:  PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,741

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer Esq., William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2460

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 32:

```

; SEQUENCE CHARACTERISTICS:

```



Query Match 5.3%; Score 316; DB 4; Length 8991;  
Best Local Similarity 21.4%; Pred. No. 8.4e-10;  
Matches 248; Conservative 138; Mismatches 479; Indels 296; Gaps 52;

Qy	6	QSPLVSSSDSP-PRPQPAFKYQF-VREPEDEEEEEEEDEDEEDLEE--LEVLERKPA	61
Db	7193	EAPAEQPKPAPAPQAPAPAPKPEKPAEQPKPEKTDDQQAEDYARRSEEEYNRLTQQQPPK	7252
Qy	62	GLSAAPVP-----TAPAAG-----APLMDF	81
Db	7253	AEKPAPAPKGTGWKQENGWYFYNTDGSMEQAGQYRAAAEGDLAAKQAELEKTEADLKKA	7312
Qy	82	GNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSVSTVPAPSPLSAAAVSPSKLPED---DE	138
Db	7313	VNEPEKPAPAPETPA--PEAPAEQPK--PAPAPQAPAPAPKPEKPAEQPKAEKTDQQAEE	7368
Qy	139	PPAR-----PPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXX	189
Db	7369	DYARRSEEEYNRLTQQQPPKAEKPAPAPKPEQAPAPAPKNSKGEQAEQYRSAAGGDLAAKQ	7428
Qy	190	XKI----MDLKEQPGNTISAGQEDFPSVLEET-AASXPSLSPLSAAFSKEHEYLGNLSTV	244
Db	7429	VELEKTEADLKK-----AVNEPEKPAPAPETPAPEAPAEQPKPAPAPQ-----	7471
Qy	245	LPTEGTLQENVSEASKEVSEKAKTLLIDRLTEFSELEYSEMGSSFSVSPKAESAVIVAN	304
Db	7472	-PAPAPKPEKPAEQPK--AEKPADQQAEDYDRRSEEEYNRL--TQQQPPKAEKPAPAPQ	7526
Qy	305	PREEI-----IVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEVVSSEKAKDSFNE	357
Db	7527	PEQPAPAPKSLKEIDESDSEDYVKEGFRAPLQSELDKQAKLSKLEEL-----SDKIDE	7580
Qy	358	KRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQ	417
Db	7581	LDAEI-AKLEKDVEDFK-----XSDGEQAGQYLAAAEDLIAKKA---ELEQ	7623
Qy	418	TNHEKDSESSNDDTSFPS---TPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTS	474
Db	7624	T--EADLKKA VNEPGKPAPAPAPE-----TPAPEAPAEQPK-----PAP	7660
Qy	475	ENXTDEKKIEEKKAIIVTEKNTSTKTSNPFFVAAQDSETDYV--TTDNLTKVTEEVVA--	530
Db	7661	ETPAPAPKPEKPAEQPKPEK-----PADQQAEDYARRSEEEYNRLTQQQPAPA	7709
Qy	531	-----NMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP	584
Db	7710	QKPEQPAKPEKPAEEPTQPEKDAEIAKLE-KNVEYFKKTDAEQTEQYLAAAEDKDLADKKA	7768
Qy	585	SFEESEA-----TPSPV-LPDI VMEAPLNSAVPSAGASVIQSSSPLEAS	628

Db 7769 ELEKTEADLKKAVNEPEKPAEETPAPAPKPEQPAEQPKPAPAPQP-APAPKPEKTTDDQQA 7827

Qy 629 SVNYESIKHEPENPPPYEEAMSVSLKVSGIKEE----IKEPENINAALQETEAPYISIIAC 684  
 :| | | | :| | | | | | :|

Db 7828 EEDYARRSEEEYNRLPQQQPPKAEKPAPAPKPEQPVPAEXPEN-----PAPAPKPAXAP 7881

Qy 685 DLIKETKLSAEPAPDFSDYSEMAKVEQPV-----PDHSELVEDSS-----PDS 727  
 :| : : :| | :| | | :| | :|

Db 7882 QPLKPEEPAEQPKPE-----KPEEPAGQPEPEKPDQQAAGEDYARRSGGEYNRFPQQ 7933

Qy 728 EPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMIEYENKEKLSALPPEGG----- 781  
 :| | | | :| | | | | | | | :|

Db 7934 QPPKAEKPAPAP----KPEQVPVAPKTL-----KKAKLAGAKSKAATKKAEL 7977

Qy 782 KPYLESFKLSLDNTKDTLLP-----DEV-----STLSKK-EKI PLQMEELSTAVYSNDD 829  
 :| | | :| | | | | | :| :| | | | :| | | :|

Db 7978 EPELEKAEAELENLLSTLDPEGKTQDELDKEAAEAELNKKVEALPNQVSELEEELSKLED 8037

Qy 830 LFISKEAQIRETETFSOSSPIEII DEFPTLISSKTD SFSKLARE----YTDLEVSHKSEI 885  
 | :| :| :| :| :| :| :| :| :|

Db 8038 NL--KDAETNNVEDY-----IKEGLEEAIATKQAELEKTPKELDAALNELGPDGDEEE 8088

Qy 886 ANAPDGAGSLPCTELPHDL SLKNIQPKVEEKI--SFSDDFSKNGSATS KVL LPPDVSAL 943  
 | :| | | :| | | :| :| :|

Db 8089 TPPPEAPAEQPKPEKPAE-ETPAPAPKPEKSADQQA EEDYARRSEEEYNRL----- 8138

Qy 944 GHTQAEIESIVKPKVLEKEAEKKLPSDTEKEDRSPSAI F SADLGKTSVVDLLYWRD---I 1000  
 | | :| | | | | :| :| | | :| :|

Db 8139 --TQQQ-----PPK-----AEKPAPAPAPKPEQAPAPAKSRGLATKKKLNLA EARI ELLL 8186

Qy 1001 KKTGV-----VFGASLFLLLS 1016  
 | | | :| | | | |

Db 8187 KKLGLEPGLKAGAGLGNLLS 8207

# RESULT 9

US-09-149-476-411

; Sequence 411, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: PZ002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311

; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,862  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,887  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,908  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/057,650

```
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
```

```
Query Match          4.8%; Score 285; DB 4; Length 92;
Best Local Similarity 57.1%; Pred. No. 9.8e-11;
Matches 52; Conservative 20; Mismatches 19; Indels 0; Gaps 0;
```

```
Qy      1079 NSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVP 1138
          |:|: |:| :| : ||||:||||||| || ||: ||||:||||:||||| : :|||
Db      2 NAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFNGITLLILAELLIFSVP 61

Qy      1139 VIYERHQAQIDHYLGLANKNVKDAMAKIQAK 1169
          ::||::: ||||:|:| | : || :|
Db      62 IVYEKYKTQIDHYVGIARDQTKSIVEKIPSK 92
```

# RESULT 10

```
US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Polypeptide
US-08-973-462-8
```

```
Query Match          4.7%; Score 276.5; DB 3; Length 1786;
Best Local Similarity 20.1%; Pred. No. 2.2e-08;
Matches 232; Conservative 201; Mismatches 471; Indels 249; Gaps 50;
```

```
Qy      30 EPEDEEEEEEEEEDEDEDLEE-----LEVLERKPAAGLSAAPVPTAPAAGAPLMDFGN 83
          | ||:| | |:|:|::|| :| |: : : : ||::
Db      228 EESVEENDEESVEENVEENVEENDDGSGVASSVEESIASSVDESIDSSIEENVAPTVE--- 284

Qy      84 DfVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPAR- 142
```

Db	285	EIVAPS-----VVESVAPSVEESVEENVEESVAENVEESVAENVEESVAENVEESVAEN 338
Qy	143	----PPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQ 198
Db	339	VEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVEESVEENVEESVAENVEESVAENVEES 398
Qy	199	PGNTISAGQEDFPSVLLETAAS-----XPSLSPLSAASFKEHEYLGNLSTVLPT-EGTL 251
Db	399	VAENV--EESVAENVEESVAENVEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVEESV 455
Qy	252	QENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVSPKAESAVI-----VANP 305
Db	456	EENVEESVAENVEESVAENVEESVAENVEESVAE-NVEESVAENVEESVAENVEESVAEN 514
Qy	306	REEIIVKNKDE----EEKLVSNNILHX-----QQELPTALTCLKVKED-----EVV 346
Db	515	VEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVEESVEENVEESVAENVEESVAENVEES 574
Qy	347	SSEKAKDSFNEKRVAVEAPMREEYA-----DFKPFERVWEVKDSKEDSDMLAAGGKIESN 401
Db	575	VAENVEESVAENVEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVEES-----VEEN 626
Qy	402	LESKVDKKCFADSLEQTNHE--KDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATE 459
Db	627	VEESV-----AENVEESVAENVEESVAENVEEIVAPTVEEIVAPTVEEIVA----PSVVE 677
Qy	460	SIATNIFPLLEDPTSENXTDE--KKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVT 517
Db	678	SVA----PSVEESVEENVEESVAENVEESVAENVEESVAENVEES---VAENVEEIVAPT 730
Qy	518	TDNLTKVT-EEVVA-NMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTS--EVMQ 573
Db	731	VEEIVAPTVEEIVAPSVVESVAPS-VEESVEENVEESVAENVE-ESVAENVEESVAENVE 788
Qy	574	ESLYPAAQ--LCPSFEESEATPSPVLPDI VMEAPLNSAVPSAGASVIQSSSPLEASSVN 631
Db	789	ESVAPTVEEIVAPSVEESVA-PS-----VEESVAENVATNLSDNLLSNLLGGIETEEIK 841
Qy	632	YESIKHEPENPPP-----YEEAMSVSLKV-SGIKEEIKE---PENINAALQETEA 677
Db	842	-DSILNEIEEVKENVVTILENVEETTAESVTTFSNILEEIQENTITNDTIEEKLLELHE 900
Qy	678	PYISIA-----CDLIKETKLSAEPAPDFSDYSEMA--KVEQPVPDHSSELVED 722
Db	901	NVLSAALENTQSEEEKKEVIDVIEEVK--EEVATTLIETVEQAEEKSANTITEIFENLEE 958
Qy	723	SSPDS-----EPVDLFSDDSI PDVPQKQDETVM LVKESLTET----- 759
Db	959	NAVESNENVAENLEKLNETVFNTVLDKVEETVEISGESLENNEMDKAFFSEIFDNVKG IQ 1018
Qy	760	-----SFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKK 810
Db	1019	ENLLTGMFRSIETSIVIQSEEKVDL-----NENVVSSILDNIENMKEGLL-NKLENISST 1072
Qy	811	E-----KIPLQME-----LSTAVYSNDDL F----- 831

Db 1073 EGVQETVTEHVEQNVYVDVDVPAMKDQFLGILNEAGGLKEMFFNLEDVFKSESDVITVEE 1132

Qy 832 -----ISKEAQIRETETETFSDSPIEIDEFPTLISSKTDSPSKLAREYTDLEVSHKSEIA 886  
: || : : || : : : | : | : : | : |

Db 1133 IKDEPVQKEVE-KETVSIIEEMEENIVDVLEEEKEDLTDKMIDAVEESIEISSDSKEETE 1191

Qy 887 NAPDGAG--SLPCTELPHD-----LSLKNIQPK-VEEKISFSDDFSKNGSATS SKVL 934  
: | : | : : : | : | : : : : : | : | : :

Db 1192 SIKDKEKDVSLVVEEVQDNDMDSEVEKVLLEKNMEEELMKDAVEIND-----ITSKLI 1244

Qy 935 LLPPDVSA LGHTQAEIESIVKPKVLEKEAEKKLPSDTE-----KEDRSPSAIFSADLGKT 989  
: : | : : : : | : | : | : : : | : | : |

Db 1245 ---EETQELNEVEAD---LIKDMEKLKELEKALSSED SKETIDAKDDTLEK VIEEEHDITT 1298

Qy 990 SVVDLLYWRDIKK 1002  
: : : : : | : : :

Db 1299 TLDEVVELKDVEE 1311

# RESULT 11

US-09-345-473E-40

; Sequence 40, Application US/09345473E

; Patent No. 6558903

; GENERAL INFORMATION:

; APPLICANT: Hodge, Martin

; TITLE OF INVENTION: No. 6558903e1 Kinases and Uses Thereof

; FILE REFERENCE: 35800/183781

; CURRENT APPLICATION NUMBER: US/09/345,473E

; CURRENT FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 40

; LENGTH: 1601

; TYPE: PRT

; ORGANISM: C. elegans

US-09-345-473E-40

Query Match 4.5%; Score 267.5; DB 4; Length 1601;

Best Local Similarity 20.5%; Pred. No. 6.8e-08;

Matches 229; Conservative 134; Mismatches 376; Indels 377; Gaps 56;

Qy 29 REPEDEEEEEEEEE-----EDEDEDLEELEVLERKPAAGLSAAPVPTAPAAGA 76  
| : || || | | | || : | || : | | : | | |

Db 462 RAKEEEERIREEAEIKEELRLRAEAKEKEKERLEK-ERLEK KAAAAAANPNP TP----- 515

Qy 77 PLMDFGNDFVPPAPRGP-----LPAAPPVAPE-----RQPSWDPS-PVS--STVPAPSP 122  
: || | | | | : : : : : | : | : : : | |

Db 516 -----IPPTPATPHSSAQQQPIPPPLSTQTS AEIQQSAQQPSVPVTMIANI PAMSP 566

Qy 123 LSA---AAVSPSK----LPEDDEPPARPPPPPPASVSPQAEPVW---TPPAPAP--AAPP 170  
|| : || : : | : | : : | | | : : ||

Db 567 TSAQPQPVLSP TSAAVPVPTTMIHVPKPSEIPVQNVATTAAPVAANNVPPSPAPFKTE DI 626

Qy 171 STPAAPKRRGSSGAVVXXXXXKIMDLKEQPGNTISAG-QEDFPSVLLETAASXPSLSPLSA 229  
|| : || : | : : : | | | | | :

Db 627 QTPTLAQ-----NTVPR TISTDASGLVINTPASIASPSPAPS 663



Qy 230 ASFKEHEYLGNLSTVLP-----TEG-----TLQENVSEASKEVSEKAKTLLIDRDLT 276  
 | : : : | | : | | | : : | :  
 Db 664 AT---DVASTTAPVTPAPTPTTTTDDGGAAAASTTTENKEEKRKSNKRKVVMEILGCDES 719  
 Qy 277 E-----FSELEYSEMGSFSVSP-----KAESAVIVANPREEIIVK 312  
 | : | : | : | : : | | :  
 Db 720 RNFALVSCRLDTSKSVTFQFAPGTDKPCITATKLLAEDCLLKVHVHIVEAQLGEVIQLI 779  
 Qy 313 NKDEEE---KLVSNNILHXQQELPTALTCLVKEDEVVSSEKAKDSFNEKRVAVEAPMRE 368  
 | | : : | | : : | : | : : | : : :  
 Db 780 NSDGKKGVGTKLAT--VLDPNSTEPPTITAVMPKD---SSAATASNTKPKIEI----- 827  
 Qy 369 EYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDSSESN 428  
 | : : | : : : : : | | | : | |  
 Db 828 -----EKTPPTRDASQE-----PNNVQVTNVRKVSQESN 856  
 Qy 429 DDTSPSTPEGIKDRSGAYITCAPFN-----PAAT---ES 460  
 : | | | | : | | : | | |  
 Db 857 AE-SVQSIP-----RPGGIIVMSPTNQDSAPPPTGAAAKPSRFQVTKSADPIATPISSS 910  
 Qy 461 IAT-NIFPILLEDPTSENXTDEKKIEEK-KAQIVTEKNTSTKTSNPFVAAQDSETDYVTT 518  
 | : | : : | | | | : | : | : | : : | : : |  
 Db 911 ISTATVIPIVA-ATPTNITSEPVIVQPITAQVITHLATPSPVSHSL---SSNSSPSATTH 966  
 Qy 519 DNLTKEVEEVANMP-EGLTDLVQEACESELNEVTGKIAIYETKMDLVQTSEVMQESLY 577  
 | : : : : : | | : | | : : | |  
 Db 967 SNMSSI--QSTTSVPGRRTVQPVSA-ESGISSSISTPHPEPT----- 1007  
 Qy 578 PAAQLCPSFEESEATPSPVLPDIVMEAPLNSAV-----PSAGASVIQSSSPLEASSVNY 632  
 | | | | | : | : | | | : : | : | :  
 Db 1008 PAITSCP-----PPVPSVPPVVSNGTLNLEVAPKQTPSATNQNVDTQHSSSTASTATL 1060  
 Qy 633 ESIKHEPEN-----PPPYEEAMSVS-----LKVSGIKEEIKEPENINAA 671  
 : | | : : : : | | : :  
 Db 1061 --VSETPATVHVTPISVPAPVQEPLVIDHHSVDLTQLDSELRKVSGVSHS-ASPSTVVES 1117  
 Qy 672 LQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELVEDSSPDSEPVD 731  
 | : | : : : | | : : :  
 Db 1118 LTSMTPTIPLACQTV-PASIGQAPAVIAAAHA-----AS 1151  
 Qy 732 LFSDDSI PDVPQKQDETVM LVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLS 791  
 | : | : | : | | : : | | | | :  
 Db 1152 LIPNASVPQSPSRLD-----AETGLAGL-----HEKLEAL-----KME 1184  
 Qy 792 LDNTKDTLLPDE---VSTLSKKEKI PLQ-----MEELSTAVYSNDDLFISKEAQIRET-- 841  
 | : | : : : | : | : : | | : : : | |  
 Db 1185 QDRRED--MGDDAIGTTTTDGDKEIPIDTLKGLAEALGKVIHADG-----RETP 1232  
 Qy 842 -----ETFSDDSPIEII-----DEFPTLISSKTDSFSKLAREYTDLEVS--HKSEIANAP 889  
 : | : : | : | : | : | | | : : | : : |  
 Db 1233 MPPDHPDLTDASTQQQLISPSNPDLVLTMTSSAVEGSASSTMI EDIDASTSAVDASMMNSMP 1292  
 Qy 890 DGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATS KVL LPPDVSALGHTQAE 949  
 | | : : : : | | : : : | | : : |  
 Db 1293 PGA-----QNSTDQIPAAMTLSMDQECAQSM TSSITR-----NTTGTKLAT 1333  
 Qy 950 IESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSAD 985

Db 1334 FENL-----ETALSSTLGTHIRQPNAPSSRD 1359

RESULT 12

US-08-978-277A-4

; Sequence 4, Application US/08978277A  
; Patent No. 6582956  
; GENERAL INFORMATION:  
; APPLICANT: Gelman, Irwin H.  
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10112-0228  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,277A  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/665,401  
; FILING DATE: 18-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Richard S  
; REGISTRATION NUMBER: 26,154  
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-408-2558  
; TELEFAX: 212-765-2519  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1596 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:

US-08-978-277A-4

Query Match 4.5%; Score 265; DB 4; Length 1596;  
Best Local Similarity 21.3%; Pred. No. 9.6e-08;  
Matches 232; Conservative 145; Mismatches 422; Indels 288; Gaps 47;

QY 2 EDLD--QSPLVSSSDSPPRPQPAFKYQFVREPEDE-----EEEEEEEEEEDEDEDLEELEV 54

Db	620	EELEKVKSATLSSTDST-----VSEMQDEVKTVGEEQKPEEPKRRVDTSVSWEA	668
Qy	55	LERKPAAGLSAAPVPTAPAAGAPL-----MDFGNDFVPPAPR-----	91
Db	669	LICVGSSKKRARKASSSDDEGGPRTLGGDSHRAEEASKDKEAGTDAVPASTQEQDQAQGS	728
Qy	92	-GPLPAAPPVAPERQPSWDPSVSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPAS	150
Db	729	SSPEPAGSPSEGEVSTWE-----SFKRLVTPRKK---SKSKLEEKAEEDSS--VEQLSTE	778
Qy	151	VSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXXKIMDLKEQPGNTISAG----	206
Db	779	IEPSREESWV-----SIKKFIPGRRKKRA-----DGKQEATVEDSGPVEI	819
Qy	207	QEDFPSVLLETAASXPSLSPLSAASFKEHEYL---GN-----LSTVLPTEGTLQENVSE	257
Db	820	NEDDPNV-----PAVVPLSEYNAVEREKMEAQGNTELPQLLGAVYVSE-----E	863
Qy	258	ASKEVSEKAKTLLID--RDLTEFSELEYSEMGSFSVSPKAESAVIVANPREEIIIVKNKD	315
Db	864	LSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVT-EPLEHTAGEAMPPEEVTEKDII	922
Qy	316	EEEKLVSNNILHXQQELPTALTKLVKE-----DEVVSSEKAKDSFNEKRVAVEAPMREEY	370
Db	923	AEETPV-----LTQTLPEGKDAHDDMTSEVDFTS-----EAVTATET	960
Qy	371	ADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDESSNDD	430
Db	961	SEALRTEEVTASGAETTDMVSAVSQLTDS-----PDTTEEATPVQEVESGVLD	1010
Qy	431	TS-----FPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPILLEDPTSENXTDEK	481
Db	1011	TEEEERQTQAILQAVADKVKEES-----QVPATQTVQRTGSKALE-----KVE	1053
Qy	482	KIEEKKAQIVTEKNTSTKTSNPPFFVA-----AQDSETDYVTTDNLTKVTEEVVANMPEGL	536
Db	1054	EVEEDSEVLASEKEKDVMPKGPVQEAEGAHLAQGSETQAT-----PESL	1098
Qy	537	TPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPV	596
Db	1099	-----EVPEVT-ADVDPVATCQVIKQLQQLMEQAVAPESS--ETLTDSETNGSTP	1144
Qy	597	LPD-----IVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEAMS	650
Db	1145	LADSDTADGTQQDETIDSQDSKATAAVRQSQVTEEEAATAQKEEPSTLPNNVPAQE----	1200
Qy	651	VSLKVSGIKKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSEMAKVE	710
Db	1201	-----EHGEEPG--RDVLEPTQQELTAAAVPVLAKTEVGQEGEVDWLD-GEKVKEE	1248
Qy	711	QPVPDHSELVEDSSPDSE-PVDLFSDDSI PDVPQKQD-ETVMLVKESLTETSFESMIEYE	768
Db	1249	QEVFVH-----SGPNSQKAADVTDSEVMGVAGCQEKESTEVQSLSLEEGEMETDVEKE	1302
Qy	769	NKE-KLSALPPEG-----GKPYLESFKLSLDNTK-----DTLLPDEVSTL	807

Db 1303 KRETKPEQVSEEGEQETAAPHEGTYGKPVLTLDMPSSSERGKALGSLGGSPSLPDQ---- 1358

Qy 808 SKKEKIPLQMEELSTAVYSNDDLFIKAEQIRETETFSDS--SPIEIIIDE--FPTLISSK 863  
| | :| :| | : :| | :| :| | :| |

Db 1359 DKAGCIEVQVQSLDITVTQTAEAV----EKVIETVVISETGESPE-ECVGAHLLPAEKSSA 1413

Qy 864 TDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFS--- 920  
| | : : :| :| :| :| :| :| :| :|

Db 1414 TGGHWTLQHAEDTVPLGPESQ-----AESIPIIVTPAPES--TLHPDLQGEISASQRE 1464

Qy 921 -----DDFSKNGSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKEAEKKLP 968  
| | :| ||| :| :| :| :| :| :|

Db 1465 RSEEDKPDAGPDADGKESTAIEKVLKAEPEILELESKSNKIVLNVIQTAVDQFARTETA 1524

Qy 969 SDTEKED 975  
:| |

Db 1525 PETHAYD 1531

# RESULT 13

5180808-2

;Patent No. 5180808

; APPLICANT: RUOSLAHTI, ERKKI I.

; TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID

;SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN

;ANTIBODIES, AND METHODS OF DETECTING THE SAME

; NUMBER OF SEQUENCES: 4

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/441,179

; FILING DATE: 27-NOV-1989

;SEQ ID NO:2:

; LENGTH: 2409

5180808-2

Query Match 4.4%; Score 263.5; DB 6; Length 2409;  
Best Local Similarity 20.1%; Pred. No. 2.1e-07;  
Matches 253; Conservative 180; Mismatches 479; Indels 345; Gaps 58;

Qy 43 EDEDEDLEELEVLER--KPAAGLSAAPVPTAPAAGAPL-MDFGNDFVPPAPRGPLPAAPP 99  
| : : | :| :| | | | | | | :|

Db 910 ENITQTSREIVISERLGPENYGAIEIRGFST---GFPLEEDFSGDF----REYSTVSH 960

Qy 100 VAPE-----RQPSWDPSPVSSSTV-----PAPSPLSAAVSPSKLP-EDDEP 139  
:| | | | | | :| :| :| :| :| :|

Db 961 IAKEETVMMEGSGDAAFRDTQTSPSTVPTSVHISHISDSEGP-SSTMVSTSAFPWEEFTS 1019

Qy 140 PARPPPPPPASVSPQAEPVWTPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQP 199  
| :| | | | | | :| :| :| :|

Db 1020 SAEGSGEQLVTVSSSVVPVL-----PS--AVQKFSGTASSII-----DEGL 1058

Qy 200 GNTISAGQEDFPSVLLETAASXPSLSPLSAAKFKEHYLGNLSTVLP----- 246  
| : : | | :| | : :| :| :| :|

Db 1059 GEVGTVNEIDRRSTILPTAEVEGTAKPVEK---EEVKVSGTVSTNFPQTIEPAKLWSRQE 1115

Qy 247 -----TEGTLQENVSEASKEVSEKAKTLLIDRDL---TEFSELEYSEMGSSE 290  
:| :| | | :| :| :| :| :| :|

Db 1116 VNPVRQEIESETTSEEQIQEEKSFESPQNSPATEQTIQFDSQFTETELKTTDYSVLTTKK 1175

Qy	291	SVSPKAESAIVANPREEIIVKNKDEEEKLVLS-----NNILHXQQELPTALTCLKVKED	343
Db	1176	TYSDDKEM-----KEEDTSLVNMSTPDPDANGLESYTTLPEATE---KSH	1217
Qy	344	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAG-GKIESNL	402
Db	1218	FFLATALVTESIPAEHVVTDSPIKKEEST-KHFPKGMRTIQESDTELLFSGGLSGSEEV	1276
Qy	403	ESKVDKKCFADSLEQTN-----HEKDSSES-SNDDTSFPSTPEGIKDRSGAYITCAPFNP	456
Db	1277	PTLPTESVNFTEVEQINNTLYPHTSQVESTSSDKIEDFNRMENVAKEVGPLVSQTDIFEG	1336
Qy	457	ATESIATNIFPLLEDPTSEN-----XTD-----EKKIEEKKAQIVTEKNTS	497
Db	1337	SGSVTSTTLIEILSDTGAEGPTVAPLPFSTDIGHQPQNQTVRWAAEIIQTSRPQTITEQDSN	1396
Qy	498	TKTSNPPFFVAAQDSETDYVTTDNLTQVTEEVVANMPEGLTPDLVQEACESELNEVTGTGI	557
Db	1397	KNSSTAEGINETTTSSDTFLARAYGFEMAKEFVTSAPK--PSDLYYEPSGEGSGEV-----	1449
Qy	558	AYETKMDLVQ-----TSEVMQESLYP-----AAQLCPSFESEATPSPVLPDIVMEAP	605
Db	1450	-----DIVDSFHTSATTQATRQESSTTFVSDGSLEKHPEVPSAKAVTADGFPTVSVMLP	1503
Qy	606	LNSAVPSAGASVIQSSSPLEASSVNYESI KHEPENPPPYEEAMSVSLKVSIGIKEEIKEP	665
Db	1504	LHSE-----QNKSSPDPTSTLS-NTVSYERSTDGSFQDRFR-EFEDSTLKPNRKKP	1552
Qy	666	ENINAALQETEAPYISIIACDLIKETK-----LSAEPAPDF-SDYSEMAKVEQPVPD	715
Db	1553	-----TE---NIIIDLKEDKDLILTITESTILEILPELTSKNTIIDIHTKPV	1599
Qy	716	HSELVE-----DSSPDSEPVDLFSDDSIPDPVQKQDETVMVLVKESLTETSFESMIE----	766
Db	1600	YEDILGMQTDIDTEVPSEPHD-SNDESNDSTQVQEIYEAANVLSLTEETFEGSADVLAS	1658
Qy	767	-----YENKEKL-----SALPPEGGKPYLESFKLSLDNTKDTLLPDEVST	806
Db	1659	YTQATHDESMTYEDRSQLDHMGFHTTGIPAP-----STETELDVLLPTATSLP	1707
Qy	807	LSKKEK--IPLQMEELSTAVYSNDDLFI SKEAQIRETETFS DSSPI----EIIDEFP TLI	860
Db	1708	IPRKSATVIP-EIEGIIKAEAKALDDMF-----ESSTLSDGQAIADQSEII---PTL-	1754
Qy	861	SSKTDSFSKFLAREYTDLE-----VSHKSEIANAPDGAGSLPCTELPH	902
Db	1755	----GQFERTQEEYEDKKHAGPSFQPEFSSGABEALVDHTPYLSIATTHLMDQSVTEVPD	1810
Qy	903	DLSLKNIQPKVEEKISFSDDFS KNGSAT-SKVL LPPDV SALGHTQ-----	947
Db	1811	VMEGSNPPYYTDTTLAVS-TFAKLSSQTPSSPLTIYSGSEASGHT EIPQPSALPGIDVGS	1869
Qy	948	-----AEIESIVKP---KVLEKEAEKKLP SDTE---KEDRSPSAIFSADLG	987
Db	1870	SVMSPQDSFKEIHVNI EATFKPSSEEY LHITEPPSLSPDTKLEPSEDDGKPELLEEMEAS	1929



Db 775 LEEKSEDSIAGSGVEHSTPDTEPGKEESWVSIKKFI PGRKKRPDQKQEAPVEDAGPTG 834

Qy 43 -EDEDED-----LEELEVLERKPAAGLSA--APVPTAPAAGAPLMDFGNDFV--PPAP 90  
::| | : :||: | | | || : | |

Db 835 ANEDSDVPVAVVPLSEYDAVEREKMEAQQAQKGAEQPEQKAATEVSKELSESQVHMMAAA 894

Qy 91 RGPLPAAPPVAPERQPSWDPSVVS---STVPAPSPLSAAAVSPSKLPEDDEPPARPPPPP 147  
| : || || : |: | | : | : : || | |

Db 895 VADGTRAATII EERSPSWISASVTEPLEQVEAEAALLTEEVLEREVIAEEEEPPTVTETPLP 954

Qy 148 PASVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXXKIMDLKEQPNTISAGQ 207  
: : | | : :

Db 955 -----ENREARGDTVVSEA 968

Qy 208 EDFPSVLLETAASXPSSLSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAK 267  
| | : || | | | : || : | : ||:

Db 969 ELTPEAV--TAA-----ETAGPLGSEEGTEASAAEETEMVSAVSQLTD 1010

Qy 268 TLLIDRDLTEFSELE-----YSEMGSFSFVSPPKAESAIVANPREEII--VKM 313  
: : | | : | : : | | : | : | :

Db 1011 SPDTTEEATPVQEVEGGVPDIEEQERRTQEVQLQAVA EKVKESQLPGTGGPEDVLQPVR 1070

Qy 314 KDEEEKLVSNNILHXQQELPTALTCLKVEDEVVSSEKAKDSFNE-----KRVAVEAPMRE 368  
: | : | : | || : || | : : | | |

Db 1071 AEAE-----RPEEQAEASGLKKETDVVLKVDAQEAKEPFTQGKVVGQTTP--- 1116

Qy 369 EYADFKPFERVWEVKDSKEDSDMLA-----AGGKIESNLESKVDDKCFADSLE-QTNH 420  
: ||: : | | | : : || | : | : : ||: | :

Db 1117 -----ESFEKAPQVTESI ESSELVTTCAETLAGVKSQ---EMVMEQAI PPDSVETPTDS 1168

Qy 421 EKB9ESSNDTSFPSTPE5X: GIKDRSGAYI-----TCAPFNPAATESIATNIFPLLE 470  
| | : | | | : | : : : | | | | : :

Db 1169 ETDGSTPVADFDPAGTTQKDEIVEI HEENEVHLVPVRGTAEAEVPAQKERPPAPSSFVFQ 1228

Qy 471 DPTSENXTDEKKIEEKKAQIVTEKNTSTKTNSPFFVAAQDSETDYVTITDNLT KVTEEVVA 530  
: | | | : | | | : | : | : | : | :

Db 1229 EETKEQSKMEDTLEH-----TDKEVSVETVSIL-----SKTE--GTQEADQYADEKTK 1274

Qy 531 NMP--EGLTPDLVQEACESELNEVTGTKIAYETKM DLVQTSEVMQESLYPAAQLCPSFEE 588  
: | || | : : || : : | : : | :

Db 1275 DVPPFFEGL-----EGSID--TGITVSREKVTEVALKGEGTEEAECKKDDALELQSH 1323

Qy 589 SEATPSPVLPDI VMEAPLNSAVPSAGASVI QPSSSPLEASSVNYESIKHEPE----- 640  
: : |||| : : : : : : | : || | : ||

Db 1324 AKSPSPVEREMVVQ-----VEREKTEAEP THVNEEKLEHELAVTVSEEVS 1369

Qy 641 -----NPPP---YEEAMSVSLKVSGIKEEIKEPENINAALQE 674  
: || | || : : |

Db 1370 KQLLQTVNVPI IDGAKEVSSLEGSPPPCLGQEEAVCTKI QV-----QS 1412

Qy 675 TEAPYISIACDLIKETKL SAEPAPDFS DYSEMAKVEQVPDPHS SELVEDSSPDSEPVDLF- 733  
: || : | : | : | : | : | | | | :

Db 1413 SEASFTLTAA--AEEEKVLGETA----NILETGETLEPAGAHLVLEEK SSEKNEDFAAHP 1466

Qy 734 SDDSIPDVPQKQDETVM LVKESLTETS FESMI EYENKEKL----- 773  
: : : | | : : : : | : | |

Db 1467 GEDAVPTGPD CQAKSTPVI VSATTKKGLSSDLEGEKTTSL KWKSDEVDEQVACQEVKVS 1526

Qy 774 --SALPPEGKPYLE--SFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDD 829  
 | | | | | | : | : | | | : | : : :  
 Db 1527 AIEDLEPENGILELETKSSKL-VQNI IQTAVDQFVRT-----EETATEMLTSE- 1573  
  
 Qy 830 LFISKEAQI-----RETETFSDDSSPIEIIDEFPTLISSKTDSDSFKLAREYTDL--E 878  
 : : | : : | | : | | : : : : : | : :  
 Db 1574 --LQTQAHVIKADSQDAGQETEKEGEEPQASAQDETPTSKEESESTAVGQAHSKSLV 1631  
  
 Qy 879 VSHKSEIANAPDGAGS-----LPCTE-----LPHD-----LSLKNI 909  
 : | | : | | | | | : | | : : : :  
 Db 1632 MSEASEKMTVEVEGSTVNDQQLLEEVLPSEEEGGGAGTKSVPEDDGHALLAERIEKSLV 1691  
  
 Qy 910 QPKVEEKISFSDDFSNGSATS KVL LPPDV SALGHTQAEIESIVKPKVLEKE----- 962  
 : | | : | | | | : | | | | | | | | hom  
 Db 1692 EPKEDEKGGDDVDDPENQNSALA-----DTDASGGLTKESPD TNGPKQKEKEDAQEVEL 1744  
  
 Qy 963 AEKKLPDTEK 973  
 | | : : : |  
 Db 1745 QEGKVHSESDK 1755

RESULT 15

US-08-994-570-5

; Sequence 5, Application US/08994570

; Patent No. 6090929

; GENERAL INFORMATION:

; APPLICANT: Scott, John D.,

; APPLICANT: Nauert, Brian J.,

; APPLICANT: Klauck, Theresa M.

; TITLE OF INVENTION: Protein Binding Domains of Gravin

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower/233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/994,570

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6090929and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27866/33451

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 5:



; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-994-570-5

Query Match 4.3%; Score 256; DB 3; Length 1780;  
Best Local Similarity 19.2%; Pred. No. 4e-07;  
Matches 221; Conservative 163; Mismatches 419; Indels 348; Gaps 45;

```
Qy      1 MEDLDQSPLVSS--SDSPPRPQPAFKYQFV-----REPEDEEEEEEEEE----- 42
      :|:  :  :  |  |  :|  :  :|  :  :|:  :|:  :|:  :|:
Db      775 LEEKSEDSIAGSGVEHSTPDTEPGKEESWVSIKKFIPGRRKKRPDGKQEQAPVEDAGPTG 834

Qy      43 -EDEDED-----LEELEVLERKPAAGLSA---APVPTAPAAGAPLMDFGNDFV--PPAP 90
      :|:  |  |  |  :  :|:  |  |  |  |  |  |  :  |  |  |
Db      835 ANEDDSDVPAVVPLSEYDAVEREKMEAQQAQKGAEQPEQKAATEVSKELSESQVHMAAAA 894

Qy      91 RGPLPAAPPVAPERQPSWDPSVPS---STVPAPSPLSAAAVSPSKLPEDDEPPARPPPPP 147
      |  :  ||  ||  :  :|  |  :  |  |  :  :  :|  ||  |
Db      895 VADGTRAATIIERSPSWISASVTEPLEQVEAEAALLTEEVLEREVIAEEEEPTVTEPLP 954

Qy      148 PASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQ 207
      :  :|  |  :|:  :
Db      955 -----ENREARGDTVVSEA 968

Qy      208 EDFPSVILLETAA SXPSLSPLSAA SFKEHEYLG NLS TVLPTEGTLQENVSEASKEVSEKAK 267
      |  |  :  ||  |  |  |  |  |  :  ||  :  |  :|  ||:
Db      969 ELTPEAV--TAA-----ETAGPLGSEEGTEASAAEETTEMVSAVSQLTLD 1010

Qy      268 TLLIDRDLTEFSELE-----YSEMGSSFSVSPKAESAVIVANPREEII--VKN 313
      :  :  |  |  :|  |  :  :  :  |  ||  :  |  :  :|  :
Db      1011 SPDTTEEATPVQEVEGGVPDIEEQERRTQEV LQAVA EKVKEESQLPGTGGPEDVLQP VQR 1070

Qy      314 KDEEEKLVSNNILHXQQELPTALT KL VKED EVVSSEKAKDSFNE-----KRVAVEAPMRE 368
      :  |  :  :  :  |  ||  :|  |  :  :  :  |  |  |  |
Db      1071 AEAE-----RPEEQAEASGLKKETDVVLKVDAQEAKTEPFTQGKVVGQTTP--- 1116

Qy      369 EYADFKPFERVWEVKDSKEDSDMLA-----AGGKIESNLESKVDKKCFADSLE-QTNH 420
      :  ||:  :|  :|  |  ||:  ||  :  |  :  :  ||:  |  :
Db      1117 -----ESFEKAPQVTESI ESSELV TTCQAETLAGVKSQ--EMVMEQAI PPDSVETPTDS 1168

Qy      421 EKDESSSNDDTSFPSTPE-----GIKDRSGAYI-----TCAPFNPAATESIATNIFPLLE 470
      |  |  :  |  |  :  :  :  :  :  |  |  |  |  :  :
Db      1169 ETDGSTP VADFDAPGTTQKDEIVEIHEENEVHLVPVRGTEAEAVPAQKERPPAPSSSFVFQ 1228

Qy      471 DPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVT TDNLT KVTEEVVA 530
      :  |  |  :  :|  |  :|  |  :  :  |  :|:  |  :  :|
Db      1229 EETKEQSKMEDTLEH-----TDKEVSVETVSIL-----SKTE--GTQEADQYADEKTK 1274

Qy      531 NMP--EGLTPDLVQEACESELNEVTG TKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEE 588
      :|  ||  |  :  :  ||  :  :  |  :  :  |  :|:
Db      1275 DVPPFFEGL-----EGSID--TGITVSREKVTEVALKGEGTEEA ECKKDDALELQSH 1323

Qy      589 SEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPE----- 640
      :  :  ||  ||  :  :|:  :  :  :  :  :  |  :  ||  |  :  :|
```

Db 1324 AKSPSPVEREMVVQ-----VEREKTEAEPHVNEEKLEHETAVTVSEEVS 1369  
 Qy 641 -----NPPP---YEEAMSVSLKVSGIKEEIKEPENINAALQE 674  
 Db 1370 KQLLQTVNVPIIDGAKEVSSLEGSPPPCLGQEEAVCTKIQV-----QS 1412  
 Qy 675 TEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQPVPDHSELVEDSSPDSEPVDLF- 733  
 Db 1413 SEASFTLTAA--AEEKVLGETA---NILETGETLEPAGAHVLVEEKSSSEKNEDFAAHP 1466  
 Qy 734 SDDSI PDVPQKQDET VMLVKESLTETSFESMIEYENKEKL----- 773  
 Db 1467 GEDAVPTGPDCAKSTPVIVSATTKKGLSSDLEGEKTTSLKWKSDDEVDEQVACQEVKVS 1526  
 Qy 774 --SALPPEGKPYLE--SFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDD 829  
 Db 1527 AIEDLEPENGILELETKSSKL-VQNI IQTAVDQFVRT-----EETATEMLTSE- 1573  
 Qy 830 LFISKEAQI-----RETETFSDSPIEIIDEFPTLISSKTDSFSKLAREYTDL--E 878  
 Db 1574 --LQTQAHVIKADSQDAGQETEKEGEEPQASAQDETPTSMAKEESESTAVGQAHSDISKD 1631  
 Qy 879 VSHKSEIANAPDGAGS-----LPCTE-----LPHD-----LSLKNI 909  
 Db 1632 MSEASEKTMTVEVEGSTVNDQQLEEVVLPSEEEGGGAGTKSVPEDDGHALLAERIEKSLV 1691  
 Qy 910 QPKVEEKISFSDDFSKNKSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKE----- 962  
 Db 1692 EPKEDEKGDDVDDPENQNSALA-----DTDASGGLTKESPD TNGPKQKEKEDAQLEVEL 1744  
 Qy 963 AEKKLPSDTEK 973  
 Db 1745 QEGKVHSESDK 1755

Search completed: January 22, 2004, 16:34:54  
 Job time : 33.5792 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:31:15 ; Search time 32.1516 Seconds  
(without alignments)  
3523.514 Million cell updates/sec

Title: US-09-830-972-29  
Perfect score: 5923  
Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	787.5	13.3	776	2 A46583	neuroendocrine-spe
2	688	11.6	208	2 I60904	neuroendocrine-spe
3	671	11.3	267	2 A60021	tropomyosin-relate
4	528	8.9	2484	2 T26216	hypothetical prote
5	519.5	8.8	2607	2 T26215	hypothetical prote
6	340.5	5.7	222	2 T26213	hypothetical prote
7	331	5.6	7962	2 I38346	elastic titin - hu
8	325.5	5.5	5327	2 T13564	microtubule-associ
9	324.5	5.5	865	2 A47282	calcium-binding pr
10	320.5	5.4	873	2 A47283	calphotin - fruit
11	299.5	5.1	3488	2 T34418	hypothetical prote
12	299	5.0	3924	2 S37431	ankyrin 2, neurona
13	290.5	4.9	1274	2 T16251	hypothetical prote

14	289	4.9	2364	2	A56577	microtubule-associ
15	289	4.9	2464	1	QRMSPI	microtubule-associ
16	287.5	4.9	971	2	T19431	hypothetical prote
17	285.5	4.8	1621	2	A82255	hypothetical prote
18	278	4.7	1948	2	S00485	gene 11-1 protein
19	274	4.6	3507	2	T34513	hypothetical prote
20	272.5	4.6	1829	2	T24583	hypothetical prote
21	272.5	4.6	2361	2	T25752	hypothetical prote
22	271	4.6	2187	2	T30826	nascent polypeptid
23	270	4.6	1230	2	T22458	hypothetical prote
24	267.5	4.5	1851	2	T19964	hypothetical prote
25	263.5	4.4	2409	1	A60979	versican precursor
26	262.5	4.4	6642	2	T29757	protein UNC-89 - C
27	261	4.4	1558	2	B71603	RESA-H3 antigen PF
28	260.5	4.4	3381	2	T42389	versican precursor
29	259.5	4.4	1828	2	A40115	microtubule-associ
30	258	4.4	1616	2	G64242	cytadherence-acces
31	257.5	4.3	5170	2	T15348	hypothetical prote
32	256	4.3	990	2	I51618	nucleolar phosphop
33	256	4.3	1634	2	T26517	hypothetical prote
34	255.5	4.3	1684	2	JW0057	gravin - human
35	253	4.3	1189	2	S56852	hypothetical prote
36	252.5	4.3	1224	2	T14007	microtubule-associ
37	252	4.3	4377	2	A55575	ankyrin 3, long sp
38	250.5	4.2	1824	1	QRHUMT	microtubule-associ
39	250.5	4.2	1830	2	A37981	microtubule-associ
40	248	4.2	1825	2	S13507	microtubule-associ
41	248	4.2	1890	2	T04556	hypothetical prote
42	248	4.2	3421	1	WZBEB6	367K tegument prot
43	246	4.2	1110	2	I51116	NF-180 - sea lampr
44	244.5	4.1	1029	2	T30351	mucin-like protein
45	243	4.1	1320	2	JC5630	TCOF1 protein - mo

# ALIGNMENTS

## RESULT 1

A46583

neuroendocrine-specific protein, splice form A - human

N;Contains: neuroendocrine-specific protein, spl|ce form B

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 05-Nov-1999

C;Accession: A46583; I60903

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: A46583

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-776 <ROE1>

A;Cross-references: GB:L10333; NID:g307306; PIDN:AAA59950.1; PID:g307307

A;Accession: I60903

A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 421-776 <ROE2>  
A;Cross-references: GB:L10334; NID:g307308; PIDN:AAA59951.1; PID:g307309  
C;Genetics:  
A;Gene: GDB:RTN1; NSP  
A;Cross-references: GDB:203968; OMIM:600865  
A;Map position: 14q21-14q22

Query Match 13.3%; Score 787.5; DB 2; Length 776;  
Best Local Similarity 31.9%; Pred. No. 3.1e-25;  
Matches 229; Conservative 91; Mismatches 190; Indels 207; Gaps 25;

```

Qy      588 ESEATPSPVLPDI--VMEAPLNSA-----VPSAGASVIQPSSSPLE---ASSVNY---- 632
      |  || | |||: :  | |:  | :  | :  ||:  | : |
Db      141 EELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKILADPLDQMKAEAYKYIDIT 200

Qy      633 --ESIKHEPENPPPYEEA-----MSVSLKVGSIKEEIKEPENINAAL-----QET 675
      | :||: :: | |:  :|| | |:| | | : : : |
Db      201 RPREEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDK- PAPVEGKI IKDHLLLEEST 259

Qy      676 EAPYISIIACDL-----IKETKLSAEPAPDFSDYSEMAKVEQPVDPHSELVED 722
      ||||  ||  :| | ||: :  : : | : ::
Db      260 FAPYID---DLSEEQRRAPQITTPVKITLTEIEPSVE-----TTTQEKTPKQDICK 309

Qy      723 SSPDSEPVDLFS---DDS-----IPDVPQKQ-----DETVMLVKESLTETSF 761
      |||: |  |||  | : |  || : : ||:
Db      310 PSPDTVPTVTVSEPEDDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEA----- 363

Qy      762 ESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELS 821
      : || | : : | :  || | | | : : :
Db      364 -KGLSYETAENPRPVGQLADRP-----EVKARSGPPTIPSPLDHEA 403

Qy      822 TAVYSNDDLFI SKEAQIRETETFS DSSPIEIIDEFP---TLISSKTDSFS----- 868
      :: | |  |||: | | : | ||
Db      404 SSAESGD-----SEIELVSEDPMAAEDALPSGYVSFGHVGGPPPPSP 444

Qy      869 -----KLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSD 921
      : || : | : | : : | : | | | | |
Db      445 ASPSIQYSILREEREAE L DSELI IESCDASSAS-----EESPKREQDSPPMKPSALD 496

Qy      922 DF-----SKNGSATS K VLL-----LPPDV SALGHTQAEIESIVKP 956
      | : | | |  |||  || | | : : |
Db      497 AIREETGVRAEERAPSRRGLAEPGSFLDYPSTEPQGPPELPPGDGAL-----EPETPMLP 551

Qy      957 KVLEKEAEKKLPSDTEKEDRSPSA-----IFSADLGKTSVVDLLYWRDIK 1001
      : | : : ||: |  : | | | : |||||
Db      552 -----RKPEEDSSSNQSPAATKGPGPLGPGAPPPLLF---LNKQKAIDLLYWRDIK 599

Qy      1002 KTG VVFGASL FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDGHPFRA 1061
      : ||: |||: | || ||| ||: ||| || ||||| | : ||: ||: |||||: |
Db      600 QTGIVFGSFLLLLFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKA 659

Qy      1062 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1121
      ||| | : | : | ||: : : || | : ||||| ||||| ||||| : |||||
Db      660 YLELEITLSQEIQKYTDCLQFYVNSTLKE LRRLFLVQDLVDLSLKFAVLMWLLTYVGALF 719

```

QY 1122 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKI QAKI PGLKRKAE 1178  
 |||||:::||:|::||:| :||| ||| : : ||||| || ||  
 Db 720 NGLTLLLMAVVSMTLPVVYVKHQAQIDQYLGLVTRTHINAVVAKI OAKI PGAKRHAE 776

## RESULT 2

I60904

neuroendocrine-specific protein C - human

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence revision 24-May-1996 #text change 05-Nov-1999

C;Accession: I60904

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: I60904

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: E

TKL

A;Residues: 1-208 <RES>

A;Cross-references: GB:L10335; NID:g307310; PIDN:AAA59952.1; PID:g307311

C;Genetics:

A;Gene: GDB:RTN1; NSP

A;Cross-references: GDB:203968; OMIM:600865

A;Map position: 14q21-14q22

Query Match 11.6%; Score 688; DB 2; Length 208;

Best Local Similarity 67.5%; Pred. No. 6.2e-22;

Matches 129; Conservative 32; Mismatches 30; Indels 0; Gaps 0;

Qy 988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKQVI 1047  
| : : ||||| : | : | : | | : | | : | | : | | : | | : |

Db 18 KSQAIDLLYWRDIKQTGIVFGSFLLLFSLTQFSVSVVAYLALAALSATISFRIYKSVL 77

Qy 1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRFLVDDLVDLSLKF 1107

Db 78 QAVQKTDEGHPFKAYLELEITLSOEIOIKYTDCLOFYVNSTLKLRLFLVODLVDSLKE 137

[illegible]

Db 138 AVLMWLLTYVGALFNGLTLLMAVVSMTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIO 197

Qy 1168 AKIPGLKRKAE 1178

[illegible]

Db 198 AKIPGAKRHAE 208

### RESULT 3

A60021

tropomyosin-related protein, neuronal - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1993 #sequence revision 03-Mar-1993 #text change 05-Nov-1999

C;Accession: A60021

R;Wieczorek, D.F.; Hughes, S.R.

Brain Res. Mol. Brain Res. 10, 33-41, 1991



Qy 39 EEEEEED-----EDEDLEELEVLERKPA-----AGLSAAP----- 67  
 :||| | : | : | : | :  
 Db 1454 NQEEEDVVAELNFHPIRQWRDEDVISLQSLKSLVAEVCITDVDASDVNEQDEESTLKIL 1513  
 Qy 68 --VPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAP--ERQPSW----- 108  
 ||: | : : || | | : | | | | : | :  
 Db 1514 KVVPSSEPSLLE--LDFTND--PKVIHVPIPLMEPATMYLEEMVEWI IADAVKEVSEMEVV 1569  
 Qy 109 -----DPSPVSSTVPAPSPLSAAAVSPSKLP-EDDEPPARPPPPPPASVSPQAEPVW 159  
 | | | | | : | | | | | | | : | :  
 Db 1570 TESEISEMAPQVSESTCPIPEPL-----ADLKLPEVEDDEKTPEPEPVVPGQVQERI IPIE 1624  
 Qy 160 TPPAPA-PAAPPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVL--LE 216  
 UE || | | | : : | : | :  
 Db 1625 VEQAPTIPQRPPRAP-----KSELPKVAKPLD 1651  
 Qy 217 TAASXPSSLPLSA-----ASFKEHEYLGNLSTVLPTTEGLQENVSEASKEVSEKAKTLL 270  
 : | : || : : : | : : | | | : : : : | : |  
 Db 1652 DSKSRVRFAPLNIKLGRTYSEEQQKELVESLERPL-TIITQQKPPEKPTEDIG--ALSPL 1708  
 Qy 271 IDRDLTEFSELEYSEMGSFSVSPKAESAVIVANPREEII VKNKDEEEKLVSNNILHXQQ 330  
 | | : | : : | : : | : | | : | |  
 Db 1709 SPNTLAEEYEEVPMMDMQS-----VPHSPQE-----KQEEIEALSEII----- 1745  
 Qy 331 ELPTALTCLKVEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFK---PFERVWEVKDSKE 387  
 | | | : : | : | : | : : | : | : : | : | :  
 Db 1746 EEPQAMKEVEKPVE-SAPEKDNESELEAPEI-INEPIRRVLVETKIMGPGKSLNEDNDDDD 1803  
 Qy 388 D-SDMLAAGGKIESNLESKVDKCFADSLEQTNHEKDSSESS---NDDTSFPSTPEGIKD 442  
 | | : | : | : : : | : : : | | | | | : | : |X50  
 Db 1804 DGSECLDSIGDLS---ERTIQR--FNTSIDDPSIRRDSFSSISSFGDRQKFRTAIENIRQ 1858  
 Qy 443 RSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSN 502  
 || : : : : | : : ||  
 Db 1859 -----DLLPFQSSVSQYLRSSPNP-----SQQLLVTN----- 1885  
 Qy 503 PFFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETK 562  
 : || : | : | | | | | : : : | : : | : : |  
 Db 1886 ----LSMDSPSD--LSPNAPPVGFENTAQFLEKLQOE-DRPSAEGSIDSSGFEEKVDHE-- 1936  
 Qy 563 MDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSS 622  
 : | | | : | | | : | | : : | |  
 Db 1937 -----GLDEFAAPP-----VHDPMQKSVFGSLGSDDMKPGS 1967  
 Qy 623 SP-----LEASSVNYESIKHEPE AMSVSLKVSGIKKEEIKEPENIN-----AAL 672  
 : | : | : : | : : | : : ||  
 Db 1968 QDDGFVFIERNEAN-----EATLKKNQKMSSHNDVIEKNYFNDNAPTAAL 2013  
 Qy 673 QETEAPYISIACDLIKETKLSAEPAPDFSDYSEMA-----KVEQPVPDHSEL 719  
 | : | | : : || : | : | | : |  
 Db 2014 --LESPIAEEARKLVQDAVESA-----SEYKKQAVDSGDEIGRELLDNVEQKIEQVKEP 2065  
 Qy 720 VEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMIEYENKEKLSALPPE 779  
 : || : | | | | : | : : | : | |  
 Db 2066 IVDSL--HKAYDGVGVFVHETV PNAVD D FVREA EKQLPESPVPEKIE-----TPE 2113



Qy 780 GGKPYLESFKLSLDNTKDTL--LPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFISKEAQ 837  
 Db 2114 -----PLVDIHDTVDKVHDEVNDFLRREPTP-PFETDDVAPLSDDKPQFGNQTP 2161  
 Qy 838 IRETETFSDDSSPIEIIDEFPTLISSKTSFSKILAREYTDLEVSHKSEIANAPDGAGSLPC 897  
 Db 2162 EEDETTFDRKGPLTIPEEVEKAAAAQNNDLD---DFDPLVTSNTGAAFGAAGVAAAA-- 2215  
 Qy 898 TELPHDLSLKNIQPKVEEKISFSDDFSKNKSATSQVLLLPDVSALGHTQAEIESIVKPK 957  
 Db 2216 -----VESLTEEM-----FGH--QKFETVPRPP 2237  
 Qy 958 VLEKEAEKKLPDTEKEDRSPSAIFSADLG-----KT-- 989  
 Db 2238 T-----PPKDISDEDVKPSTV---NLGPSHHSHSPSSPHHSILKHHGDAWIDFKTVP 2286  
 Qy 990 -SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVQI 1048  
 Db 2287 PCVLDVIYWRDAKSAIVLSLALLVLFVLAKYPLLTVVVTSLLLALGAAAGFRVFKKVEA 2346  
 Qy 1049 AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFA 1108  
 Db 2347 QIKKTDSEHPFSEILAQDLTLQEKVHAQADVFEHATCIANKLKKLVFVESPLESIKFG 2406  
 Qy 1109 VLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA 1168  
 Db 2407 LVLWSLTYYIASWFSGFTLAILGLLGVSFVPKVYESNQEAIDPHLATISGHLKNVQNIIDE 2466  
 Qy 1169 KIPGLK 1174  
 Db 2467 KLPFLR 2472

# RESULT 5

T26215

hypothetical protein W06A7.3a - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T26215

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26215

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-2607 <WIL>

A;Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a

A;Experimental source: clone W06A7

C;Genetics:

A;Gene: CESP:W06A7.3a

A;Map position: 5

A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2; 2586/2

Query Match 8.8%; Score 519.5; DB 2; Length 2607;  
 Best Local Similarity 20.4%; Pred. No. 1.8e-13;  
 Matches 292; Conservative 175; Mismatches 474; Indels 491; Gaps 55;

Qy	1	MEDLDQSPVLVSSSDSPRPQPAFKYQFVREPEDEEE-----EE	38
Db	1397	LEKVEVQP--DLSQNSPAPHKIIDLHF-NIPKDHEDYGNDYVFPFGTESSEESQKADGNQE	1453
Qy	39	EEEEED-----EDEDLEELEVLERKPA-----AGLSAAP-----	67
Db	1454	NQEEEDVVAELNFHPIRQWRDEDVISLQSLKSLVAEVGCITDVDASDVNEQDEESTLKIL	1513
Qy	68	--VPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAP--ERQPSW-----	108
Db	1514	KVVPSEPSLLE--LDFTND--PKVIHVPIPLMEPATMYLEEMVEWIIADAVKEVSEMEVV	1569
Qy	109	-----DPSPVSSTVPAPSPLSAAAVSPSKLP-EDDEPPARPPPPPPASVSPQAEVW	159
Db	1570	TESEISEMAPQVSESTCPIPEPL-----ADLKLPEVDEDEKTPEPEPVVPGVQVQERIIPIE	1624
Qy	160	TPPAPA-PAAPPSTP-----AAPKRRGSSGAVVXXXXXKIMDLKEQPGNTISAGQE---	208
Db	1625	VEQAPTIPQRP RPAPKSELPKVAKPLDDSKS-----RVRFAPLNIKLGRTYSEEQQKEL	1678
Qy	209	----DFPSVLLETAASXP-----SLSPLSAASFKEHEY--LGNLSTVLPTEGTLQEN	254
Db	1679	VESLERPLTII-TQQKPPEKPTEDIGALSPLSPNTLAEYEEVPMMDMQSVPHSPQEKQEE	1737
Qy	255	VSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVSPKAESAVIVANPREEIIIVK--	312
Db	1738	I-EALSEIIEEPQAM-----KEVEKPVESAPEKDNESLEAPEIINEPIRRVLVETK	1787
Qy	313	-----NKDEEEKLVSNNILHXQQELPTALTCLKVEDEVVSSEKAKDSFNEKRVAVEA	364
Db	1788	IMGPGKSLNEDNDDDDDGSECLDSIGDL-----SERTIQRFN---TSIDD	1829
Qy	365	P--MREEYADFKPFERVVEVKDSKED--SDMLAAGGKIESNLESKVDDKKCFADSLEQTNH	420
Db	1830	PSIRRDSFSSISSFGDRQKFRTAIENIRQDLLFPQSSVSQYLRSSPNP---SQQLLVTNL	1886
Qy	421	EKDSESSNDDTSFPSTPEGI-----KDRSGA--YITCAPFNPAATESIATNIF	466
Db	1887	SMDSPS---DLSPNAPPVGFENTAQFLEKLQQEDRPSAEGSIDSSGFVKVDHEGLDEF	1943
Qy	467	PLLEDPTSEN-----XTDEKK-----IEEKKAQIVTEKNTSTKTSNPPFFVAAQDS	511
Db	1944	PPVHDPMQKSVFGLSGSDDMKPGSQDDGFVFIERNEANEATLKKNQKMSSHNDVIEKNY	2003
Qy	512	ETDYVTTDNLTKVTEEVANMPEGLCESELNEV-TGTKIAYE----TKMDv	566n
Db	2004	FNDNAPT---AALLESPIAEEARKLVQDAVESASEYKKQAVDSGDEIGRELLDNVEQKIE	2060
Qy	567	QTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLE	626
Db	2061	QVKEPIVDSLHKAYDGVGDFVH-ETVPAV-DDFVREAE-----	2097
Qy	627	ASSVNYESIKHEPENPPPYEEAMSVSLKVSGIKEEIKEPENINAALQETEAPYISIIACDL	686
Db	2098	-----KQLPESPVP-----EKIETPE-----PLVDIHDTV	2122

Qy 687 IK-----ETKLSAEPAPDFSDYSEMAKVEQVPDHSSELVEDSSPDSEPVDLFSDDSI PDV 741  
 | : | | | : | : | : | :  
 Db 2123 DKVHDEVNDNFLRREPTPPFE-----TDDVAPLSDDKPQFGNQ T----- 2160

Qy 742 PQKQDETVM LVKESLTETS FESMI EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLP 801  
 : | | | | : | : | : | : | : | :  
 Db 2161 -PEEDETTFDRKGPLT-----IP 2177

Qy 802 DEVSTLSKKEKIPLQMEELSTAVYSNDDL FISK EAQIRETETFS DSSPIEIID EFP TLI S 861  
 : | | : | : | : | : | : | : | :  
 Db 2178 EEV-----EKA AAAQNND-----LDDFDPLVT 2199

Qy 862 SKT-----DSFSKLAR-----EYTDLEVSHKSEIANAPD 890  
 | | : | : | : | : | : | : | :  
 Db 2200 SNTGA AFGAAVGAAA VESLTEEEMFGHQKFETVPRPPTPKDISDEDVK-PSTVNLGPS 2258

Qy 891 GAGSLPCTELPHD-----LSLKNIQPKVEEKIS-----FSDDFS 924  
 | | : | : | : | : | : | : | :  
 Db 2259 HHHSHPS--PHHSILKHHGDAWIDFKTVPPCAQNAFSPGEIMFLLAFFVYLSCFASFFS 2316

Qy 925 KNGSATSKVLLL-----PPDVSA LG----- 944  
 | : : | : | : | : | : | : | :  
 Db 2317 KSLPLLDNLLSLVVYLSISLIHVKHHRKFRWNEEQATTMSKLGAVGRGLYALIAFIVNI 2376

Qy 945 --HTQAEIESIVKPKVLEKEAEKKLPSTDEKEDRSPSAIF SADLGKTSVVDLLYWRDIKK 1002  
 : : | | | | | : | | | : | : | : | :  
 Db 2377 VLRVGLNVALVVGVA VSAHEAYKLTKS-----SGVLRKKEVL DVIYWRDAKK 2423

Qy 1003 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY 1062  
 : : | : | : | : : : | | | | : : | : | : | : | : | :  
 Db 2424 SAIVLSLALLVLFVLAKYPLLT VVTYSLLLALGAAAGFRVFKKVEAQIKKTDSEHPFSEI 2483

Qy 1063 LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFN 1122  
 | : : : | : : : | : : | : : : | : : : | : : | : : : | : : : | : :  
 Db 2484 LAQDLTLPQEKVHAQADVFEHATCIANKLKKLVFVESPLESIKFGLVLSLTYIASWFS 2543

Qy 1123 GLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLK 1174  
 | | | | : : | | : | : | : : | : | : | : | : | : | :  
 Db 2544 GFTLAILGLLG VFSVPKVYESNQEAIDPHLATISGHLKNVQNIIDEKLPFLR 2595

# RESULT 6

T26213

hypothetical protein W06A7.3b - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T26213

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26213

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-222 <WIL>

A;Cross-references: EMBL:Z78066; PIDN:CAB01523.1; GSPDB:GN00023; CESP:W06A7.3b

A;Experimental source: clone W06A7

C;Genetics:

Query Match 5.7%; Score 340.5; DB 2; Length 222;  
Best Local Similarity 32.4%; Pred. No. 1.4e-07;  
Matches 66; Conservative 52; Mismatches 83; Indels 3; Gaps 2;

```

RESULT 7
I38346
elastic titin - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C;Accession: I38346
R;Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A;Title: Titins: giant proteins in charge of muscle ultrastructure and
elasticity.
A;Reference number: A57430; MUID:96026330; PMID:7569978
A;Accession: I38346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-7962 <RES>
A;Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g101742
C;Genetics:
A;Gene: GDB:TTN
A;Cross-references: GDB:127867; OMIM:188840
A;Map position:      -2q31      ----

```

```

Query Match          5.6%; Score 331; DB 2; Length 7962;
Best Local Similarity 20.0%; Pred. No. 4.7e-05;
Matches 235; Conservative 144; Mismatches 404; Indels 392; Gaps 48;

Qy          32 EEEEEEEEEEEDEDEDLEEELE--VLERK---PAAGLSAAPV-----PTAPA 73
              |:| ||||| | | | :|| | |:| | : : | | | | |
Db          6176 EEEREEEEEEAEVTEYEVMEEEPYYVVEEKLHIISKRVEAEPAEVTERQEKKIVLKP KIPA 6235

Qy          74 AGAPLMDFGNDFVPPAPRGPLPAAP-PVAPERQPSWDPSPVSSTVPAPSPLSAAAVSPSK 132
              : | :| || : ||: : |||| | | | |
Db          6236 K-----IEEPPPAKVPEAPKKIIVEKK-----VPAPVP-KKEKVPPPK 6272

```

Qy 133 LPEDDEPPA---RPPP-----PPPASVS-----PQAEPVWTPPAPAPAAPPS-TPAA 175  
 :||: : | : || | || : | | | | : |  
 Db 6273 VPEEPKKPVPEKKVPPKVIKMEELPAKVTEKHMQITQEEKVLVAVTKKEAPPKARVPEE 6332  
  
 Qy 176 PKRRGSSGAVVXXXXKIMDLK----EQPGNTISAGQEDF---PSVLLETAASXPS-LSPL 227  
 ||| |:: || |:: : : | : | : :  
 Db 6333 PKR-----AVPEEKVLKLKPKREEEPKAKVTEFRKRVVKEEKVSIKAPKREPQPIKEV 6385  
  
 Qy 228 SAASFKEHEYLGNLSTVLPTEGTLQEN-VSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286  
 : || | ||: | || : | : : : | || | : |  
 Db 6386 TIMEEKERAY-----TLEEEAVSVQREEEYEEYE---EYDYKEFEFEYEPTTEE 6429  
  
 Qy 287 GSSFS-----VSPKAESAVIV--ANPREEI 309  
 : POTE ||| | : | : :  
 Db 6430 YDQYEEYEEREYERYEEHEEYITEPEKPIPVKPVPEEPVPTKPKAPPAKVLKKAVPEEKV 6489  
  
 Qy 310 IV----KNKDEEEKLVSNILHXQQELPTALTCLKVEDEVVSSEKAKDSFNEKRVAVEAP 365  
 | | | : : : : : | : | | : | | |  
 Db 6490 PVPIPKKLKPPPKVPEEPKKVFEEKIHISITK--REKEQVTEPAAKVPMKPKRVVAEEK 6547  
  
 Qy 366 MREEYADFKPFERVWEVKDSKEDSMDLAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSE 425  
 : : | || | : : | : : : : | : ||: |  
 Db 6548 VPPVPRKEVAPPVRVPEVPKELEPEEV----AFEEEVVTHVE-EYLVEEEEEEYIHEEE-E 6600  
  
 Qy 426 SSNDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDEKKIEE 485  
 : : | | : : | | : : | | : : || ||  
 Db 6601 FITEEEVVPVIPVKVPE-----VPRKPVPEEKKPVVPVKKKEAPPKVPPEVPKKPEE 6652  
  
 Qy 486 KKAQIVTEKNTSTKTSNPPF-----VAAQDSETDYVTTDNL----- 521  
 | : : | |Pp\*950XA| :  
 Db 6653 KVPVLI PKKEKPPPAKVPEVPKKPVPEEKVPVPVPPKVEAPPKVPPEVPKKPVPEKKVPV 6712  
  
 Qy 522 -----TKVTEEVVANMPEGLTPDLVQACESELNEVTGTKIAYETKMDLVQTSEV 571  
 || | : || | | : | : : | : | ||  
 Db 6713 PAPKKVEAPPKVPPEVPKKLIPEEKKPTVPVKKVEAPPPKVPKKREPVPVPVALPQEEEV 6772  
  
 Qy 572 M-QESLYPAAQLCPSFESEATP----- 593  
 : : | : : | || | |  
 Db 6773 LFEEIIVPEEEVLP--EEEEVLPEEEVLPEEEVLPEEEIIPPEEEVPPEEEYVPEEE 6830  
  
 Qy 594 -----SPVLPDIVMEAPLNSAVPSAGASVIQ-----PSSSPLEASSVNYESI- 635  
 |||: : | : || | : | : | : |  
 Db 6831 EFVPEEEVLPEVKPVVPAPVPEIKKKVTEKKVVI PKKEEAPPKVPPEVPKKVVEEKRII 6890  
  
 Qy 836 --KHE-----PENPPPYEEAMSVy EI---KEPENIN 669  
 | | || || || ||: : || |  
 Db 6891 LPKEEEVLPEVTEEPPEEPISEEEIPEEPPSIEEV-----EEVAPPRVPEVIK 6939  
  
 Qy 670 AALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELVEDSSPDSEP 729  
 | : | | : | | || : : | : : |  
 Db 6940 KAVPEAPTVP-----PKKVEAPP-----AKVSKKIPEEKVPVPVQKKEAPP 6980  
  
 Qy 730 VDLFSDDISIPDVQKQDETVMVLKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFK 789  
 : : ||: | : || | | : || | :  
 Db 6981 A-----KVPEVPKKVPEKKVLV-----PKKEAVPPAKGR----- 7009

```

Qy      790 LSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQIRETETFS DSSP 849
          | : | : : | | : : | : : | | : : | : : | : :
Db      7010 -----TVLEEKVSVAFRQEVVVKERLELEVVEAEVEE--IPEEEEFHEVEEYFEEGE 7059

Qy      850 IEIIDEFPTLISSKTDSPFSKLAREYTDLEV--SHKSEIANAPDGAGSLPCTELPHDLSLK 907
          : : | | | : : : : : : | | : : | : | : : | |
Db      7060 FHEVEEFIKLEQHRVEEEHVRVEKVRVIEVFEAEEVEVFEEKPKAPPKGP-----EISEK 7113

Qy      908 NIQPK-----VEEKISFSDDFS KNGSATS KVL L LPPDV SAL 943
          | | | | | : : : : : : : : : :
Db      7114 IIPPKKPPTKVVPKRKEPPAKVPEVPKKIVVEEKVRVPEE-----PRVPP----- 7157

Qy      944 GHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSP 978
          : : : : | | | | : : | : : |
Db      7158 ----TKVPEVLPPK--EVVPEKKVPVPPAKKPEAP 7 86

```

RESULT 8

T13564

microtubule-associated protein homolog - fruit fly (*Drosophila melanogaster*)

N;Alternate names: hypothetical protein EG:49E4.1

C;Species: *Drosophila melanogaster*

C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C;Accession: T13564

R;Spanos, L.; Papagiannakis, G.; Siden-Kiamos, I.; Louis, C.

submitted to the EMBL Data Library, April 1999

A;Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.

A;Reference number: Z17689

A;Accession: T13564

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-5327 <SPA>

A;Cross-references: EMBL:AL031128; PIDN:CAA20006.1

C;Genetics:

A;Cross-references: FlyBase:FBgn0025392

A;Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1

A;Note: EG:49E4.1

C;Superfamily: *Drosophila* 576K microtubule-associated protein homolog

Query Match 5.5%; Score 325.5; DB 2; Length 5327;

Best Local Similarity 23.6%; Pred. No. 4.6e-05;

Matches 275; Conservative 151; Mismatches 520; Indels 219; Gaps 53;

```

Qy      28 VREPEDEEEEEEEEE---DEDEDLEELEVLERKPAAGLS-----AAPVPTAPAAG 75
          : : : | : | | : : | : : | : : | : : | : :
Db      1277 MEQVKDKEEHEQKIESGIITEKEAKKSASTPEEKETSDITSDDDELPAQLADPTTVPPKSA 1336

Qy      76 APLMDFGND FVPPA-----PRGPLPAAPPV-APERQPSWDPSV SSVTPAPSPLSAA 126
          | | : | | : | | | | : | | : | |
Db      1337 KDREDTGSIESPPTIEEAIEVEVQAKQEAQKVPAPAEAAIKTEKSPLASK-ETSRPESAT 1395

Qy      127 AVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPA-----APK 177
          : | | | | : | : : | : | : | : |
Db      1396 ----GSVKEDTEQTKSKKSPVPSRPESEAKDKKSPFASGEASRPESVAESVKDEAGKA 1451

Qy      178 RRGSSGAVVXXXXKIMDLKEQPGNTISAGQE-----DFPSVLLETAASXP-SLSPLSA 229
          | | | : | | | : | | | | | : |

```

Db 1452 RRESIAKTHKDESSLDKAKEQESRRESIAESIKPESGIDEKSALASKEASRPESVTDKS- 1510

Qy 230 ASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSS 289  
 || : : | | | : || | : : | | | :

Db 1511 ---KEPSRRESIAESLKAESTKDEKSAPPSKEASRPGSVVESVKDETEKSKEPSRRESIA 1567

Qy 290 FSVSPKAESAVIVANPREEIIIVKNKDEEEKLVSN---ILHXQQELPTALTCLVKEDEVV 346  
 | | | : | | : || | | : : | : : | : :

Db 1568 ESAKPPIEFRE-VSRP-ESVIDGIKDESAKPESRRDSPLASKEASRPESVLESVKDEPIK 1625

Qy 347 SSEKAK-----DSFNEKRVAVE-APMREEYADFKPFERVWEVKDS--KEDSDMLAAGGKI 398  
 | : || : : : || : : : : | | | : | | | :

Db 1626 STEKSRRRESVAESFKADSTKDEKSPLTSKDIS-RPESAVENVMDAPFKETSRPESAVGSM 1684

Qy 399 ESNLESKVDDKKCFADSLQVTHPEP7KDESLZSNDDTSFP-STPEGIKDRSGAYIT( 455  
 : || : | | : : : || | | | | | | | |

Db 1685 KDESMSK-----EPSRRESVKDGAAQSRETSRPASVAESAkd--GADDLKELSRP 1732

Qy 456 AATE-----SIATNIFLLEDPTS-----ENXTD--EKKIEEKKAQIVTEKN--TSTK 499  
 : | || || : | : | || || : : | || : | :

Db 1733 ESTTQSKEAGSIKDEKSPLASEEASRPASVAESVKDEAEKSKEESRRESVAEKSPLPSKE 1792

Qy 500 TSNPFFVAAQDSETDYVTTDNLTkvTEEVVANMP----EGLTPDLVQEACESELNEVTGT 555  
 | | || : | | | | | | | | | : : |

Db 1793 ASRPASVA--ESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKDE----AE 1845

Qy 556 KIAYETKMDLVQTSEVM--QESLYPAAQLCPSFEESEATPSPVLPDIVME-APLNSAVPS 612  
 | | : : | : : || : : : : | | : || | |

Db 1846 KSKEESRRESVAEKSPLPSKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEAS 1905

Qy 613 SPLEASSVNYESIKHEPENPPPYEAMSSVSLKVSGIKEEI---KEPENIN 669  
 ||| : | | : : | | : : || : | |

Db 1906 RPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKDEAEKSKEESRRE 1965

Qy 670 AALQETEAPYISIACDLIKETKLSAEPAPPDFSDYSEMAKVEQPVPDHSELVEDSSPDSEP 729  
 : : : | || | | | : : | | | | : | |

Db 1966 SVAEKSPLP-----SKEASRPASVAESIKDEAEKSKEE---SRRESVAEKSP---- 2009

Qy 730 VDLFSDDSIPDVPQKQDETVMVLESILTETSFESMIE--YENKEKLSALP-PEGGKP--Y 784  
 : | : | || : : : : | | : : | | | | :

Db 2010 -----LPSKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASV 2058

Qy 785 LESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEE----LSTAVYSNDDLFIKSEAQIRE 840  
 || | : : | : | | || || : | | | : || |

Db 2059 AESIKDEAEKSK-----EESRRESAAEKSPLPSKEASRPASVAESVKDEADKSKEESRR- 2112

Qy 841 TETFSDDSSPIEII--DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCT 898  
 | : : : | : | : : | : : | | | : ||| :

Db 2113 -ESMAESGKAQSIKGDQSPLKEVSRPES---VAESVKDDPVKSK-EPSRRESVAGSVTAD 2167

Qy 899 ELPHDLSLKNIQPKVEEK-----ISFSDDFSKNGSATSkvLLLLPPDVSAIGHQTQAE 949  
 | : : | : | | | : | | : : :

Db 2168 -----SARDDQSPLESKGASRPESVVDsvKDEAEKQES-----RRESK 2205

Qy 950 IESIVKPKVLEKEAEKKL-PSDTEKEDR-SPSAIFSADLGKTSVVDLLYW---RDI 1000  
 || : : | : : : : || : : | : : : ||

Db 2206 TESVIPPkAKDDKSPKEVLQPVSMETETIREDADQPMKPSQAESRRESIAESIKASSPRDE 2265

```

Qy      1001 KKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDGHPFR 1060
          |          | |          | : : | : : | | | | |
Db      2266 KSP-----LASKEASRPGSV-----AESIKYDLDPQIIKDDKSTE-HSRR 2305

Qy      1061 AYLESEVAI-SEELVQKYSNSALGH 1084
          | | : | : | | : : | |
Db      2306 ESLEDKSAVTSEKSVSRPLSVASDH 2330

```

# RESULT 9

A47282

calcium-binding protein calphotin - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C;Accession: A47282

R;Martin, J.H.; Benzer, S.; Rudnicka, M.; Miller, C.A.

Proc. Natl. Acad. Sci. U.S.A. 90, 1531-1535, 1993

A;Title: Calphotin: a *Drosophila* photoreceptor cell calcium-binding protein.

A;Reference number: A47282; MUID:93165729; PMID:8094559

A;Accession: A471;

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-865 <MAR>

A;Cross-references: GB:L02111; NID:g157031; PIDN:AAA28405.1; PID:g157032

A;Experimental source: photoreceptor cells

A;Note: sequence extracted from NCBI backbone (NCBIN:124955, NCBIP:124956)

C;Genetics:

A;Gene: FlyBase:Cpn

A;Cross-references: FlyBase:FBgn0010218

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; von Willebrand factor type C repeat homology

C;Keywords: calcium binding

```

Query Match          5.5%; Score 324.5; DB 2; Length 865;
Best Local Similarity 22.3%; Pred. No. 4e-06;
Matches 229; Conservative 128; Mismatches 375; Indels 297; Gaps 49;

```

```

Qy      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVP 118
          | : ||| | | ||: | ||| | : | | | : | | : |
Db      9 PVSAPVAAPV-TPSAVAAPVQVVSPPAAVAPAP APIAVTPVAPPPTLASVQPATV--TIP 65

Qy     119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAP---APAAPP--ST 172
          ||:|::||:|:| | | | |||: || : || | | || || | : |
Db     66 APAPIAAASVTP---VASVAPPVVAAPTTPAA-SPVSTPVAVAQIPVAVSAPVAPPVAAT 121

Qy     173 PAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAASF 232
          | : : | : | | | | | : | | | : | | : | : |
Db     122 PTPVVQIPVAAPVIAT-----PPVAASA-----PT----PAAVTPVISPVIAS-- 160

Qy     233 KEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSV 292
          | : | | | | | : : | : : | : | : | : |
Db     161 -----PPVVPANTT----VPVAAPVAAPVAAPVVPVLAP-----AV 194

Qy     293 SPKAESAVIVAN-----PREEIIVKNKDEEEKLVSNILHXQQELPTALTCLVKEDEVVS 347
          : | | : || | || | | : | : | 2 || :
Db     195 APAV--APVVAETPAPPPVAEIPVAT-----IPECVAPLIPEVSVVA 234

```



Qy 348 SEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVD 407  
 : | : | | | | | : : : : |  
 Db 235 T---KPLAAAEPVVVAPPATET-----PVVAPAAASPHVSVAPAVETAVVAPVS 280

Qy 408 KKCFAADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFP 467  
 || : : | : | || ||  
 Db 281 -----ASTEPPV---AAATLTAPETPAL----- 301

Qy 468 LLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEE 527  
 | : | | || | | | |  
 Db 302 -----APVVAESQ-----VAA-----NTVVATPP 320

Qy 528 VVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPA----- 579  
 | || : | : | : | || : | || : | || : ||  
 Db 321 TPAPEPETIAPPVV-----AETPEVASVAVA-ETTPPVV--PPVAAESI-PAPVVATTPV 371

Qy 580 -AQLCPSFEESEATPSPVLPDIMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHE 638  
 | | : : | : | || : : : | || : | | : | :  
 Db 372 PATLAVTDPDVTASAVPELPPVIAPSPVPSAVAETPVDLAPPVLPVAAEPVPAVVAEET 431

Qy 639 PENPPPYEEAMSV-SLKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEP 697  
 || | | : : : | : : | | : | || : | : : | ||  
 Db 432 PETPAPASAPVTIAALDIPEVAPVIAAPSDAPAEAPSAAAPIVSTPPTTASVPETTAPPA 491

Qy 698 P-----DFSDYSEMAKVEQVPDPHSEL-----VEDSSPDSEPVDLFSDDSI--DVPQK 744  
 | | || | : | || : | | : | || : | : |  
 Db 492 AVPTEPIDVSVLSE-AAIETPVAPPVEVTTTEVAVADVAPPEAAADLIEPVEPPAPIPDL 550

Qy 745 QDETVMLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEV 804  
 : : | : : | : : p : : || : : : : \$ 60  
 Db 551 LEQTTSVPAVEAAESTSSPIPE-----TSLPPPNEAVASPEVAVAPITAPEPIPEPEP 603

Qy 805 STLSKKEKIPLQ---MEELSTAVYSNDDLFIKSEAQIRETETFSOSSPIEIIIDEFPTLI 860  
 | : | || : : : : | || | : || | | | || :  
 Db 604 SLATPTEPIPVEAPVVIQEAVDV-----EVPVTETST---SIP-ETTVEFPEAV 649

Qy 861 SSKTDSFSLKAREYTDLEV-SHKSEIANAPDGAGSLPCTELP-----HDL 906  
 : | | : | : : : || || | || : | : :  
 Db 650 AEKV-----LDPAITEAPVTTQEPDVANINDGA---PATEITTPAVEIVTAAAEVSDIAI 701

Qy 907 KNIQPKVEEKISFSDDFSKNGSATSQVLL---LP-----PDVSALG----- 944  
 | | | : : : : : : : : : | : ||  
 Db 702 PVIDPPVPQEIAVA-EIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAAEVP 760

Qy 945 -----HTQAEIESIVKPKVLEKEAEKKLPSDTEKE Q8ANFSADLGKTSVVDLLY 996  
 : | | : | | : || : : : || | : | : ||  
 Db 761 ITAGDNPDNSTVGISEVV-PTIAEKPVEEVPTSEIPEQSSSPSD--SVPVAK--ITPLL- 814

Qy 997 WRDIKKTGV 1005  
 || : | |  
 Db 815 -RDLQTTDV 822

C;Species: Drosophila melanogaster  
C;Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000  
C;Accession: A47283  
R;Ballinger, D.G.; Xue, N.; Harshman, K.D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 1536-1540, 1993  
A;Title: A Drosophila photoreceptor cell-specific protein, calphotin, binds calcium and contains a leucine zipper.  
A;Reference number: A47283; MUID:93165730; PMID:8434015  
A;Contents: photoreceptor cells  
A;Accession: A47283  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-873 <BAL>  
A;Cross-references: GB:L05080; NID:g157071; PIDN:AAA28420.1; PID:g157072  
A;Note: sequence extracted from NCBI backbone (NCBIN:124958, NCBI:124959)  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; von Willebrand factor type C repeat homology

Query Match 5.4%; Score 320.5; DB 2; Le  
Best Local Similarity 22.5%; Pred. No. 6e-06;  
Matches 235; Conservative 122; Mismatches 367; Indels 321; Gaps 50;

```

Qy      59 PAAGLSAAPV-PTAPAAGAPLMDFGNDFVPPAPRGPLPAAP----PVAPERQPSW-DPSP 112
      | :  ||| | :| ||  ::  | ||  | ||| |  ||| |  :|  |
Db      9 PVSAPVAAPVTPSAVAAPVQVVSPAAVAVAPAVVAPAPAAPTAVTPVAP--PPTLASVQP 66
                        DEGQ 2330
Qy     113 VSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPP-----A 163
      : ||||| :||| :|  ||  | |||  | || |||
Db     67 ATVTVPAPAPIAAASVAP---VASVAPPVVAAPTPPA-----ASPVSTPPVAVAQIPVAV 118

Qy     164 PAPAAPP----STPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLET-- 217
      || |||  || ||  ::|  | |
Db     119 SAPVAPPVAATPTPVAP-----IPVAAPVIATPPVAASAPT 154

Qy     218 -AASXPSLSPLSAAASFKEHEYLGNLSTV:      ENVSEASKEVSEKAKTLLIDRDLT 276
      || | :|| : |  :| |  | | :  : |  :: |
Db     155 PAAVTPVVSPIAT-----PPVVPANTT---VPVAAPVAAPVAAPVVPVLA 199

Qy     277 EFSELEYSEMGSFSVSPKAESAVIVAN-----PREEIIVKNKDEEEKLVSNILHXQQE 331
      :| :|  | :||  | || |
Db     200 P-----AVAPAV--APVVAETPAPPPVAEIPVAT----- 226

Qy     332 LPTALTKLKLVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDM 391
      :| : | : | || : |  : | | |  |  |
Db     227 IPECVAPLIPEVSVVAT---KPLAAAEPVVVAPPATET-----PVVAPAAASPH 272

Qy     392 LAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCA 451
      ::  :| : : : |  ||  :  : | :| |
Db     273 VSVAPAVETAVVAPVS-----ASTEPPV---AAATLTTA 302

Qy     452 PFNPAATESIATNIFLLEDPTSENXTDEKKEEKKAQIVTEKNTSTKTSNPFVAAQDS 511
      | ||  | :| |  |||
Db     304 PETPAL-----APVVAESQ-----VAA--- 320

Qy     512 ETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEV 571
      |  |  | || : | :|  :| ||  :| ||  :| |
Db     321 -----NTVVATPPTPAPEPETIAPPVV-----AETPEVASVAVA-ETTPPVV--PPV 364

```

Qy 572 MQESLYPA-----AQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPS 622  
 ||: || | : : | : || : : | : || : |  
 Db 365 AAESI-PAPVVATTPVPATLAVTDPDVTASAVPELPPVIAAPSPVPSAVAETPVDLAPPVL 423  
  
 Qy 623 SPLEASSVNYESIKHEPENPPPYEEAMSV-SLKVSGIKKEIKEPENINAALQETEAPYIS 681  
 | : | | : || | | : : : | : : | | : | | : |  
 Db 424 PPVAAEPVPAVVAEETPETPAPASAPVTIAALDIPEVAPVIAAPSDAPAEAPSAAPIVS 483  
  
 Qy 682 IACDLIKETKLSAEPAP-----DFSDYSEMAKVEQVPDPHSEL-----VEDSSPDSEPV 730  
 : : | || | | | | : | || | : | | : |  
 Db 484 TPPTTASVPETTAPPAAVPTEPIDVSVLSE-AAIETPVAPPVEVTTEVAVADVAPPEAAE 542  
  
 Qy 731 DLFSDDSIIP--DVPQKQDETVMLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESF 788  
 || : | : | : : | : : | : : | : : ||  
 Db 3 DLIIIEPVEPPAPIPDLLEQTTSVPAVEAAESTSSPIPE-----TSLTARNENALQPEV 595  
  
 Qy 789 KLSLDNTKDTLLPDEVSTLSKKEKIPLQ---MEELSTAVYSNDDLFIKSEAQIRETETF 844  
 : : : : | | : | || : : : | || | : | ||  
 Db 596 AVAPITAPEPIPEPEPSLATPTEPIPVEAPVVIQEAVDV-----EVPVTETST- 644  
  
 Qy 845 SDSSPIEIIDFPTLISSKTD SFSKLAREYTDLEV-SHKSEIANAPDGAGSLPCTELP-- 901  
 | | | ||| : : | | : | : : : || || | || :  
 Db 645 --SIP-ETTVEFPEAVA EKV-----LDPAITEAPVTTQEPDVANINDGA---PATEITTP 693  
  
 Qy 902 -----HDSLKNIQPKVEEKISFSDDFSNGSATSQVLL---LP-----PDVS 941  
 | : : | | | : : : : : : : : : : | : ||  
 Db 694 AVEIVTAAAEVSDTAIPLIDPPVPQEIAVA-EIPETETKPAEVIVEQSTIPIEAPVPEVS 752  
  
 Qy 942 ALG-----HTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSA 980  
 : | | : | | : || | : | : : ||  
 Db 753 KYAEPVISEAPAAEVPITAGDNPNTSVGISEVV-PTIAEKPVEEVPTSEIPEQSSSPSD 811  
  
 Qy 981 IFSADLGKTSVVDLLYWRDIKKTGV 1005  
 | : | : || || : |  
 Db 812 --SVPVAK--ITPLL--RDLQTTDV 830

RESULT 11

T34418

hypothetical protein F12F3.3 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T34418

R;Fulton, B.; Wohldmann, P.

submitted to the EMBL Data Library, July 1998

A;Description: The sequence of *C. elegans* cosmid F12F3.

A;Reference number: Z21521

A;Accession: T34418

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-3488 <FUL>

A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3

A;Experimental source: strain Bristol N2; clone F12F3

C;Genetics:

A;Gene: CESP:F12F3.3

A;Map position: 5

A; Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 5.1%; Score 299.5; DB 2; Length 3488;  
Best Local Similarity 20.6%; Pred. No. 0.0003;  
Matches 272; Conservative 187; Mismatches 475; Indels 387; Gaps 57;

```
Qy      1 MEDLDQSPLV--SSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEV--- 54
      ::::| | ::      | : :| | : ::| | | :      | | : : |::|
Db      340 VDEVDDSTVLEEKDDGDDKSKPKTKKKIIKKKETPESEQVTAAEPEQQKISEVDVQSV 399

Qy      55 -----LERKPAA---GLSAAPVPT-----APAAGA-----PLMD-----FGNDF 85
      :|| | | || | :      | |      | |      | |
Db      400 ETEVGAKKKPDAEKPTDLKAKKDSKSKKSDEPEASTEESTTEKPTNDKTSKKSAEKKT 459

Qy      86 VPPAPR---GPLPAAPPVAPERQPSWDPSPVSSTVPA-----PSPLSAAAV 128
      | |      || | || : : | |      | : ||      | : : :
Db      460 VKPKKEVTGKPLEAKKPVEDKKDASQPSSSKESSPPTDGKKKKQIPKALFIPDEISSRFG 519

Qy      129 SPSKL-----PEDDEPPARPP--PPPPASVS-----PQAEPVWTPPA
      || :      :      | : | | |||      : || : :
Db      520 DPSTMHSETNITTTIRGREGSADAKTPLVEPLSASVSMKVFTLVESAKEKAEFSEFKRRSE 579
                                           164 50XE

Qy      165 AP-----AAPPSTPAAPK-----RRGSSGAVVXXXXKIMD----LKEQPGN----- 201
      |      || : : | : : | : : | : | : : || :
Db      580 TPDDKSRKKEGLPPAKKSEKKEVTAEKQSTEALIESKKKEVDESKISEQQPSDKNKSEV 639

Qy      202 ----TISAGQEDFPSVLLETAASXPSLSPLSAAASFKEHEYLGNLSTVL---PTEGTLQEN 254
      :|| | |      | : : | : : | : ||      : : :
Db      640 VGVPEKAAGPETKKDV--SEIEEVPKKKTIKKKTEKSDSSISQKSNVLKPADDDKSKSDD 697

Qy      255 VSEASKEVSEKAKTLLID-----RDLTEFSELE-----YSEMGSSFSVSP 294
      | : | | : : | : |      | : | |      : | || : |
Db      698 VTDKSKKTTEDQTKVATDSKLEKAADTTKQIETETVVDDKSKKKVLKKKTEKSDSF-ISQ 756

Qy      295 KAESAVIV--ANPRE----EIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKED----- 343
      | : : : | | : : | | : : | | : : : : | | : |
Db      757 KSETPPVVEPTKPAESEAQKIAEVNKAQKQKEVDDNL---KREA EVAAKKIADEKLKIEA 813

Qy      344 -----EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAG 395
      || : : | | : : : : | : : | : : | : |
Db      814 EANI KKTAEVEAAKKQKEKDEQLKLETEVVSKKSAAEKLELEKQAQIKKAAEAD---AVK 870

Qy      396 GKIESNLESKVD--KKCFAD--SLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCA 451
      : | | : : | : || || || : : | : | : | | : : |
Db      871 KQKELNEKNKLEAAKSAADKLKLEEESAASKKVSSESVKF---GEEKKTKAGEKTVQV 927

G      Qy      452 PFNPAATESIATNIFPLL      SENXTDEKKIEEKKAQIVTEKN$B$DKTSNPFVAAQDS 511
      | : : : | | | | | : ||| : || ||| : | : : |
Db      928 ESEPTSKKTIDTKDVGATE-PADE--TPKKKIIKKK---TEKSDSS-----ISQKS 972

Qy      512 ETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEVTGK-----IAYETKMDL 565
      || : : : | | : : | | || : : : | | || : : |
Db      973 ATD---SEKVSQKEQDEPTKPAVSETQMVTADKSKKQKETDEKLKLD AEIAAKTKQEA 1029

Qy      566 VQTSEV-MQESLYPAAQLCPSFEESEATPSPVLPD-IVMEAPLNSAVPSAGASVIQPSSS 623
      : | : : | | : : | : | | : : | : : | | : : :
Db      1030 DEKSKLDAQEKIKKVS-----EDDAARKEKELNDKLKLESEIATKKASADKLKLEEQAQ 1083
```

Qy 624 PLEASSVNYESIKHEPENPPPYE-EAMSVSLKVSGIKEEIKEPENINAALQETEAPYISI 682  
 :|: | |: | :| : : : | | : | |:: | | :  
 Db 1084 AKKAAEV--EAAKKQKEKDEQLKLDTEAASKKAAAEKLELEKQAQIKKA---AGADAVKK 1138

Qy 683 ACDLIKETKLSAEPAPDFS-----DYSEMAKVEQVPDPHSEL-----VE 721  
 :| :| :| | : : | | :| | : :| |  
 Db 1139 QKELDEKNKLEANKKSAAGKLIKIEESAAKSKQTVVEEQAKLDAQTKAKTAEKQTKLEKDE 1198

Qy 722 DSSPDSEPVDLF-----SDDSI-----PDVP 742  
 |: :|| : : || ||  
 Db 1199 KSTKESESKETVDEKPKKKVLKKKTEKSDSSISQKSETSKTVVESAGPSESETQKVADAA 1258

Qy 743 QKQDET-----VMLVKESLTETSFSMIEYENKEKLSA-----LPP 778  
 :|| || : |:| | :| :| |:| | :| A |  
 Db 1259 RKQKETDEKQKLEAEITAKKSADE---KSKLEAESKLKAAEVEAAKKQKEKDEQLKLD 1315

Qy 779 EGGKPYLESFKLSLDNTKDTLLPDEVSTLSKK-----EKIPLQME 818  
 | : || |: || : |: ||: |: :  
 Db 1316 EAASKKAAAEKLELEKQSHIKKAAEVDVKKQKELEEKQRLESEEAATKKADA EKLKLEE 1375

Qy 819 ELSTAVYS-----NDDLFI SKEAQIRETETFS---DSSPIEIIIDEFP 857  
 : | : : | | | |:| | : :| | :|| |  
 Db 1376 KKKAAEIALIEIQKEQEKLAEQSRLEDEAKKSAEKQKLESETKSKQTEEAPKESVDEKP 1435

Qy 858 --TLISSKTDSFSKLAREYTDLEVSHKSEIANAP-DGAGSLPCTELPHDLSLKNIQPKVE 914  
 : : || | :| :| ||: | : | | :| | :| | ||  
 Db 1436 KKKVLKKKT-----EKSDSSISQKSKSAKSTVDAAETL---ESDFNLVEKKTQKVE 1484

Qy 915 -----EKISFSDDFSKNGSATS KVL LPPDV SALGHTQAEIESIVK 955  
 | || : : | | | TRTS 1901  
 Db 1485 QSPDESTSATIKRDPAQKTEEISKQDDGDEKKT TTDGKPPKPEDSEA-----TPKKRVVK 1539

Qy 956 PK-----VLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVD 993  
 | | || :| | ||| | | | | :  
 Db 1540 KKTQKSDSVASDASLADVSKLSDDVVEKPKKKVLKKKTEKSDSVISETSSVDTIKPESVE 1599

Qy 994 L 994  
 :  
 Db 1600 I 1600

RESULT 12

S37431

ankyrin 2, neuronal long splice form - human

N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid ankyrin

N;Contains: ankyrin 2, short form

C;Species: Homo sapiens (man)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999

C;Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569

R;Chan, W.

submitted to the EMBL Data Library, September 1993

A;Reference number: S37431

A c S37431.425Xt

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-3924 <CHA>  
 A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288  
 R;Otto, E.; Kunitomo, M.; McLaughlin, T.; Bennett, V.  
 J. Cell Biol. 114, 241-253, 1991  
 A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins  
 reveal a family of alternatively spliced genes.  
 A;Reference number: A39643; MUID:91302466; PMID:1830053  
 A;Accession: A39643  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-2077 <OT1>  
 A;Cross-references: GB:X56957  
 A;Accession: B39643  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1443,3585-3924 <OTT>  
 A;Cross-references: EMBL:X56958  
 R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux,  
 S.E.; Ward, D.C.; Forget, B.G.  
 Genomics 10, 858-866, 1991  
 A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin  
 gene.  
 A;Reference number: A40334; MUID:92009921; PMID:1833308  
 A;Accession: A40334  
 A;Molecule type: DNA  
 A;Residues: 463-474,'PE',477-495 <TSE>  
 A;Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648  
 R;Chan, W.; Kordeli, E.; Bennett, V.  
 J. Cell Biol. 123, 1463-1473, 1993  
 A;Title: 440-kD ankyrinB: structure of the major developmentally regulated  
 domain and selective localization in unmyelinated axons.  
 A;Reference number: A49462; MUID:94075409; PMID:8253844  
 A;Accession: A49462  
 A;Status: preliminary; nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-3924 <RES>  
 A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288  
 C;Genetics:  
 A;Gene: GDB:ANK2  
 A;Cross-references: GDB:127607; OMIM:106410  
 A;Map position: 4q25-4q27  
 C;Superfamily: ankyrin; ankyrin repeat homology  
 C;Keywords: alternative splicing  
 F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>  
 F;2-11-3924/Product: ankyrin 2, short form #status predicted <MA2on50Xm  
 F;63-95/Domain: ankyrin repeat homology <AN01>  
 F;96-128/Domain: ankyrin repeat homology <AN02>  
 F;129-161/Domain: ankyrin repeat homology <AN03>  
 F;162-190/Domain: ankyrin repeat homology <AN04>  
 F;191-223/Domain: ankyrin repeat homology <AN05>  
 F;232-264/Domain: ankyrin repeat homology <AN06>  
 F;265-297/Domain: ankyrin repeat homology <AN07>  
 F;298-330/Domain: ankyrin repeat homology <AN08>  
 F;331-363/Domain: ankyrin repeat homology <AN09>  
 F;364-396/Domain: ankyrin repeat homology <AN10>  
 F;397-429/Domain: ankyrin repeat homology <AN11>  
 F;430-462/Domain: ankyrin repeat homology <AN12>

F;463-495/Domain: ankyrin repeat homology <AN13>  
 F;496-528/Domain: ankyrin repeat homology <AN14>  
 F;529-561/Domain: ankyrin repeat homology <AN15>  
 F;562-594/Domain: ankyrin repeat homology <AN16>  
 F;595-627/Domain: ankyrin repeat homology <AN17>  
 F;628-660/Domain: ankyrin repeat homology <AN18>  
 F;661-693/Domain: ankyrin repeat homology <AN19>  
 F;694-726/Domain: ankyrin repeat homology <AN20>  
 F;727-759/Domain: ankyrin repeat homology <AN21>  
 F;760-792/Domain: ankyrin repeat homology <AN22>  
 F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 5.0%; Score 299; DB 2; Length 3924;  
 Best Local Similarity 21.5%; Pred. No. 0.00037;  
 Matches 258; Conservative 167; Mismatches 417; Indels 358; Gaps 62;

Qy 14 DSPP-----RPQPAFKYQFVREPEDEEEEEEEEEDEDEDLE-----ELEVLERK 58  
 | | : | : : : : | : | | |  
 Db 1648 DIPDETQSTQKQHKPSLGIKKPVRRLKEKQKQKEEGLQASAEKAEKKGSSEESLGED 1707

Qy 59 PAAGLSAAPVPTAPAAGAPLMD-----FGNDFVPPAPRG----- 92  
 | ||: ||: | : ||: | : | : |  
 Db 1708 P--GLAPEPLPTVKAT-SPLIEETPIGSIKDKVKALQKRVEDEQKGRSKLPPIRVKKGEDV 1764

Qy 93 -----PLPAA-PPVAPERQPSWDSP-----VSSTVPAPSPL 123  
 | || | : || | || | || | :  
 Db 1765 PKKTTHRPHPAASPSLKSERHAPGSPSPKTERHSTLSSSAKTERHPPVSPSSKTEKHSPV 1824

Qy 124 SAAA-----VSPSKLPEDDEPPA-----RPPPPPPASVSPQAEPVW 159  
 | : | || | | : | | | : : | |  
 Db 1825 SPSAKTERHSPASSSSKTEKHSPVSPSTKTERHSPVSSTKTERHPPVSPSGKTDKRPV- 1883

Qy 160 TPPAPAPAAPPSTPAAPKRR---GSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLE 216  
 : | || : | : | || | || : | : |  
 Db 1884 SPSGRTEKHPPVSPGRTEKRLPVSPSGRTD-----KHQPVSTAGKTEKHLVPVSPSG 1934

Qy 217 TAASXPSSLSPLSAAS-FKEHEYLGNLSTVLPT-----EGTLQENVSEASKEVSEKAKT 268  
 | : || | : | : | : | : | : | |  
 Db 1935 KTEKQPPVSPTSSTERIEETMSVRELKAFQSGQDPSKHKTGLFEHKSAKQKQPQEKGV 1994

Qy 269 -----LLIDRDLTEFSELEYSEMSSFSVSPKAES--AVIVANPREEIIIVKNKDE-- 316  
 : | | : : | : | | | | | : | | : |  
 Db 1995 RVEKEKGPILTQRE-AQKTENQTIKRGQRLPVTGTAEKRGVVRVSS----IGVKKEDAAG 2049

Qy 317 -EEKLVSNNI-----LHXQQELPTALTKLVKE-----DEVV 346  
 : || : : | : | : | : | | :  
 Db 2050 GKEKVLSHKIPEPVQSVPEEESHRESEVPK--EKMADEQGMDLQISPDRKTSTDFSEVI 2107

Qy 7 SSE-KAKDSFNEKRVAVEAPMREEYAD---FKPFERVVEVKDSKED-----+0--- 388  
 | : | : : | : | | : : | :  
 Db 2108 KQELEDNDKYQQFRLSEETEKALHLDQVLTSPFNTTFPLDYMKDEFLPALSLQSGALDG 2167

Qy 389 -SDMLAAGGKIESNLESKVDKKCFADSLEQTNHE--KDESSNDDTSFPSTPEGIKDRS 444  
 | : | | | : | | | : | : | : : :  
 Db 2168 SSES LKNEGVAGSPCGSLMEGTPQISSEESYKHEGLAETPETSPELSLF--SPKKSEEQT 2225

Qy 445 GAYITCAPFNPAATESIATNIFLLEDPTSENXTDEKKIEEKKAIQIVTEKNTSTKTSNPF 504

```

      |      :      | |      | ||:: |      ||: | : : |||::
Db      2226 G-----ETKESTKTETTTEIRSEKEHPTTKDITGGS--EERGATVTVTEDSETSTESFQKE 2277

Qy      505 FVAAQDSETDYVTTDNLTQVTEEVVA-NMPEGLTPDLVQEACESELNEVTGTKIAYETKM 563
      :|      |: |      :|      |||      :||      |      |: |:|
Db      2278 ATLGSPKDTSPKRQDDCTGSCSVALAKETPTGLT----EAAACDEGQRTFGSS-AHKT-- 2330

Qy      564 DLVQTSEVMQESLYPAAQLCPSEFESEATPSPVL-----PDIVMEAP--LNSAV 610
      ||      |||      : :||: | |      | |: :| | |:
Db      2331 ---QTDSEAQES-----TATSDETKALPLPEASVKTDGTESKPGQGVIRSPQGLELAL 2380

Qy      611 PSAGASVIQ-----PSSSPLEASSVNYESIKHE-PEN--PPPYEEA-MSVSLKVSG 657
      || : |:      |||| | : : |: |: | |: |: | |: |
Db      2381 PSRDSEVLSAVADDSLAVSHKDSLEASPVLEDNSSHKTTPDSLEPSPLKESPCRDSLESSP 2440

Qy      658 IKEEIKE---PEN--INAALQETE--APYISIACDLIKETKLSAEPAPDFSDYSEMAKVE 710
      :: ::|      | : : ||: :||      |: |::      || | :| :|
Db      2441 VEPKMKAGIFPSHFPLPAAVAKTELLTEVASVRSRLLRD-----PDGS--AEDDSLE 2490

Qy      711 QPVPDHSELVEDS-----SPDSEPVDLFSDDSI PDVPQKQDETVMLVKESLTETSFESMI 765
      |      : |:| |      |||: : | : |      : : | |
Db      2491 Q-----TSLMESSGKSPLSPDTPSSEEVSYEVTPKTTDVSTPKPAVIHECAEED----- 2539

Qy      766 EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKK---EKIPLQMEELS 821
      : || || | |      | ||: || : ||: :|: :| | | |
Db      2540 DSENGEKKRFTPEE-----EMFKMV---TKIKMF-DELEQEAQKRDYKKEPKQEES 2589

Qy      822 TAVYSNDDLFI SKEA-QIRETETFS DSSPIEI IDEFPTLISSKTD SFSKLAREYTDLEVS 880
      : | : | | : : : | : | | |      | |:|:: | :
Db      2590 S---SDPDADCSVDVDEPKHTGSGEDES GSV-----PVLVTSES RKVSSSSSES----- 2633

Qy      881 HKSEIANAPDGAGS-----LPCTELPHDLSLKN-----IQPKVEEKISF--SDD 922
      : |:| || |      | : || : :      || | : : | :|
Db      2634 -EPELAQLKKGADSGLLPEPVIRVQPPSPLPSSMDSNSSPEEVQFQPVVSKQYTFKM NED 2692

Qy      923 FSKNGSATSKVLLLPPDV SALGHTQAEIESIVKPKVLEKEAEKKLPSD-----TEKEDRS 977
      : : :      ||::| | |      || |||
Db      2693 TQEEPGKSEE-----EKDSESH LAEDRHAVSTEAEDRS 2725

```

RESULT 13

T16251

hypothetical protein F35A5.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000

C;Accession: T16251

R;Leimbach, D.

submitted to the EMBL Data Library, January 1996

A;Description: The sequence of *C. elegans* cosmid F35A5.

A;Reference number: Z18485

A;Accession: T16251

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1274 <LEI>

A;Cross-references: EMBL:U46675; NID:g1166613; PID:g1166621; PIDN:AAB52641.1;

GSPDB:GN00028; CESP:F35A5.1

A;Experimental source: strain Bristol N2; clone F35A5



C;Genetics:

A;Gene: CESP:F35A5.1

A;Map position: X

A;Introns: 1272/2

Query Match 4.9%; Score 290.5; DB 2; Length 1274;

Best Local Similarity 20.5%; Pred. No. 0.00017;

Matches 221; Conservative 126; Mismatches 417; Indels 313; Gaps 45;

```
Qy      13 SDSPPRP--QPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPAAGLSAAPVPT 70
      | :|| | || :| | |||| | | : || :|| : ||
Db      2 SRAPPTPIKNPAKKWKPPWESVDEEEEME-----VDEETPAPSKLEKKPSLKRKDAPTKP 56

Qy      71 APAAGAPLMDFGNDVFPAPRGPPLAAPPVAPERQPSWDPSPVSSTVPAPSPLSAAAVSP 130
      | : || | | : | | : | | : || |
Db     57 VPSPGAP-----SPVPIKNPVKKWKAPWEDDEPMEEAPAAP-----VP 94

Qy     131 SKLPEDDEP---PARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRGSSGAVVX 187
      : | | | | : | : | : | | | | :
Db     95 AKKVRDPSPKKVPAPKPRDASPKKIMAAKK---EPETLPAVPP-TPVKNPVKKFKAPWED 149

Qy     188 XXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXP-SLSPLSAASFKEHEYLGNLSTVLP 246
      : | : | | : | : | | : | | | | : : |
Db    150 DEVDVEDVKDAP--TVPAKKTPLKKKEPAAAAKPRDPSPKKAAPSKEHDPI-----VPP 202

Qy     247 TEGTLQENVSEASKEVSEKAKTLLIDRDL--TEFSELE-----YSEMGSFSVSPKA 296
      | | : | | : | : | | : | | | :
Db    203 T-----PIKNPAKKWKPPWEDDEVPTTEEIKEPEPATRKVPALKKKKEPSTSVKPV 252

Qy     297 ESAVIVANPREEIIVKNKDE-----EEKLVSNNI LHXQQELPTALTKL 339
      : : | : : | | : | | : : : | :
Db    253 D-----PSPTKKVPVKKEPEVPPTPIKNPTKKWKPPWEDETPVEEV--KEPPVPEKKAPV 305

Qy     340 VKEDEVVSSEKAKDSFNEKR-----VAVEAPMREEYADFKPFERVWEVKDSKEDS 389
      : | : : | : | | : | : | : | | | :
Db    306 LKKKDPAPAAKARDPSPSKAAPKKVEPSSPVVPTPVKNPVKKYKP---PWEVDDEPAE- 361

Qy     390 DMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYIT 449
      | | : || | : | | : : | | : | :
Db    362 -----EVKKPSAPEKK--TPVLKRKEPEPSSTTPSSDPSPKKAAPAVKPRDSSPKK 410

Qy     450 CAPF--NPAATESIATNIFPLLEDPTSE-----NXTDEKKIEE-KKAQIVTEKNTSTKTS 501
      | : | | | : : | : | | : || | : : |
Db    411 ATPLQADPKAQEVPPTPV----KNPVKKYKPPWEVDDEDPVEEVKQPEAPAKKTPVLKRK 466

Qy     502 NPFFVAAQDSETDYVTTDNLTKVTEEVVANMP-----GLTPDLVQ----- 542
      | || : | | | | | | |
Db    467 EP---AAKD-----TAKPATSKTPET---PEKKDPVKPRDSSPKKVAAPKPSAQAPAT 513

Qy     543 -----EACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEES 589
      | | : : | : | : : | | | | :
Db    514 PVKNPVKKWRPPWEDDETADDVSKPTDAKKTPSLAKKDPAPAKESLKPADTKAPAKPR 573

Qy     590 EATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAM 649
      : : | | | | : | | | | : | : | :
Db    574 DPSPKKVAP-----TAPEKKTPLAKKEPAGPADSKTKEPEKSKPRDPSPKKAVP 623
```

```

Qy      650 S-----VSLKVSGIK--EEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDFSD 702
      :      :      : : | | | : : :      : | :      : | |
Db      624 AKPVPKTEVAPAAVKKPEPISKPKDTAPKKAEPNSPVV-----PPTP----- 665

Qy      703 YSEMAKVEQVPVDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFE 762
      | : | |      | |      : : | | |      | : : : | : | | : | |
Db      666 -----VKNPVKWKWKPPWEDDDAPAKPVSL-----PEPEKKTPVLAKKAPTKPDSE 710

Qy      763 SMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELST 822
      : | | |
Db      711 -----AAADPVSGP----- 719

Qy      823 AVYSNDDLFISKEAQIRETETFSDDSPIEII-----DEFPTLISSKTD SFSKLAREYT 875
      | : | : : | : : : | | : : : | : : : |
Db      720 ---SSKDPKLAKKAPVKP----RDPSPMKAVPIKPAPKTEVPPAVVKKPEPVAK----- 766

Qy      876 DLEVSHKSEIANAPDGAGSLPCTELPHDL SLKNIQPKVE-----EKISFSDDFS KN GSA 929
      : | | | : : : | : : : | : : : | : : : |
Db      767 SRDPSPKK--AKAEPNSPVVPPT--PVKNPVKKWKPPWEDDDAPAEPVNVPEPEKKTPVL 822

Qy      930 TSKVLLLPPDVS-----ALGHTQAEIESIVKPKVLEKEAEKKLP S DTEKEDRSP 978
      | : | | | | | : : | : : | | | : | | |
Db      823 AKKTPVKPRDPSPKKAVPAKPSTKTDAPPVSVKKPEPVSKPKE---PSPKKAEPNSP 876

```

#### RESULT 14

A56577

microtubule-associated protein MAP 1B - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Feb-1997

C;Accession: A56577

R;Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.

Eur. J. Cell Biol. 57, 66-74, 1992

A;Title: Identification of two distinct microtubule binding domains on recombinant rat MAP 1B.

A;Reference number: A56577; MUID:92347374; PMID:1639092

A;Accession: A56577

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2364 <ZAU>

A;Cross-references: GB:X60550

A;Experimental source: brain

A;Note: nucleotide sequence not given; conceptual translation not complete

C;Superfamily: microtubule-associated protein MAP1B

```

Query Match          4.9%;  Score 289;  DB 2;  Length 2364;
Best Local Similarity 20.9%;  Pred. No. 0.00047;
Matches 275;  Conservative 158;  Mismatches 462;  Indels 422;  Gaps 61;

```

```

Qy      30 EPEDEEEEEEEEEDEDEDLEELEVLERKPAAGLSAAPVPTAPAAGAPLMDFGNDFVPPA 89
      | | | | | | | : : | | : : | | | | : | | :
Db      909 EAEQSEEEGEEEEEDKAEDAREEDHEPDKTEAEDYVMAVVDKAAEAGVTEDQY--DFL--- 963

Qy      90 PRGSESAAPP--VAPEEQPSWDPSVSSVTPAPSPLSAAAVSPSKLPEDDEPPARPP *p975X9
      | | | : | : : | | : | | | | : | | |
Db      964 ---GTPAKQPGVQSPSREPA--SSIHDETLPGGSESEATA-----SDEENREDQPEEF 1011

```

Qy 148 PASVSPQAEPVWTPPPAPAPAAPPSTPAAPKRRGSSGAVVXXXXXKIMDLKEQPGNTISAGQ 207  
 | : : | | || : | : | | |  
 Db 1012 TATSGYTQSTIEISSEPTPMDEMSTP-----RDVMTDETNNETESPSQ 1055  
  
 Qy 208 E-----DFPSVLLETAASXP---SLSPLSAASFKE---HEYLGNLSTVLPTEGTLQENV 255  
 | : | | | : || | : : | : || : | : :  
 Db 1056 EFVNITKYESSLYSQEYSKPVVASFNGLSDGSKTDATDGRDYNASASTISPPSSMEEDKF 1115  
  
 Qy 256 SEAS-----KEVSEK-----AKTLLIDRDLTE 277  
 | : : | : || : | | | : | :  
 Db 1116 SKSALRDAYRPEETDVKTGAELDIKDVSDERLSPAKSPSLSPSPSPSPPIEKTPLGERSV-N 1174  
  
 Qy 278 FS---ELEYSEMGSFS-VSPKAESAVI---VANPRE---EII VKNK----- 314  
 || : : | : : || | : | | : : :  
 Db 1175 FSLTPNEIKASAEGEATAVSPGVTQAVVEEHCASPEEKTLEVVSQSVTGSAGHTPY 1234  
  
 Qy 315 ---DEEEKLVSNILHXQQELPTALTCLKVKEDEVVSSE--KAKDSFNEKRVAVEAPMRE 368  
 || : : : | : | | | : || | | : : || |  
 Db 1235 QSPTDEKSSHLPTTEVTENAQAVP-----VSFEFTEAKDE-NER--SSISPMDE 1279  
  
 Qy 369 EYADFK-PFERVWE-----VKDSKED---SDMLAAGGKIESNLESKVDDKKCFADSL 415  
 | : | | : : : | : | | : | | | : | : |  
 Db 1280 PVPDESPIEKVLSPLRSPPLIGSESAYEDFLSADDKALGRRSESPFEGKNGKQGFSD-- 1337  
  
 Qy 416 EQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNP--AATESIATNIFPILLED- 471  
 | : | | | : : : | : | : | : | : | : |  
 Db 1338 -----KESPVS-DLTSPLYQDKQEEKRAGFIPIKEDFSPEKKASDAEIMSSQSALALDE 1390  
  
 Qy 472 -----PTSENXTDEKKIEBKQAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLT 523  
 || : : : | : : | | : | : | : | : :  
 Db 1391 RKLGGDGSPTQVDVSQFGSKEDTKMSISEGTVSDKSATPVDEGAEDT---YSHMEGVAS 1447  
  
 Qy 524 VTEEVVA--NMPEGLTPDLVQACESELNEVTGTKIAYETKMDLVQTSEVMQES-LYPAA 580  
 | : || : || | | : : | : : : || | | : : | :  
 Db 1448 VSTASVATSSFPPTTDD-VSPSLHAEVGSPPHSTEVDDSLSVSVVQTPTTFQETEMSPSK 1506  
  
 Qy 581 QLCP-----SFESEATPSPVLPDIVMEAPLNSAV-----P 611  
 : || : : : | : : | : | : |  
 Db 1507 EECPRMSPSPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSLAMDFSRQSPDHP 1566  
  
 Qy 612 SAGASVIQ-----PSSSPLEASSVNYESIKHE--PENPPPY-----EEAMSVS-LKVSG 657  
 : || : : | : : | : | : | : | : | : | : |  
 Db 1567 TVGAGMLHITENGTEVDYSPSDIQDSSLSHKIPPTTEPSYTDNDLSELISVSQVEASP 1626  
  
 Qy 658 IKKEIKEPENINAALQETE-----APYISIACDLIKE---TKLSAEP----- 696  
 | | : || : | | : : : || :  
 Db 1627 STSSAHTPSQIASPLQEDTLSDVVPVRDMSLYASLASEKVQSLEGEKLSPKSDISPLTPR 1686  
  
 Qy 697 -----APDFSDYSEMAK-----VEQPVPDHSELVEDS----- 723  
 : | || : || : | : | : : | :  
 Db 1687 ESSPTYSPGFSDDSTSGAKESTAAYQTSSSPPIDAAAAEPYGRSSMLFDTMQHHLALSRLD 1746  
  
 Qy 724 -----SPDSEPVDLFSDD---SIPDVP-----QKQD 746  
 || | | | : || : :  
 Db 1747 LTTSSVEKDNGGKTPGDFNYAYQKPESTTESPDEEDYDYESHEKTIQAHDVGGYYEYKTE 1806  
  
 Qy 747 ETVML-----VKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKD 797

```

      | :      | | : | :      | | : | | | |      : :
Db      1807 RTIKSPCDSGYSYETIEKTTKTPEDGGYSCEITEKTTRTPEEGGYSY-----EISEK 1858

Qy      798 TLLPDEVS--TLSKKEKIPLQMEELSTAVYSNDD-----LFISKEA 836
      | | | | | | | : : : | : | :
Db      1859 TTRTPEVSGYTYEKTERRRLLDDISNGYDDTEDGGHTLGDCSYSYETTEKITSFPESES 1918

Qy      837 QIRETETFSDDSP-----IEIIDEFPTLISSKTD SFSKLAREYT-----DL 877
      | | | : | | | : : : : | | : | | | |
Db      1919 YSYETTTKTTRSPDTSAYCYETMEKITKTPQASTYSYETSDRCYTPERKSPSEARQDVDL 1978

Qy      878 -----EVSH-KSEIANA---PDG---AGSLPCTELPHDLSLKNIQP-----KV 913
      | | | : : : | : | | | | : | :
Db      1979 CLVSSCEFKHPKTELSPSFINPNLEWFAGEEPTESERPLTQSGGAPPPSGGKQQQQC 2

Qy      914 EEKISFSDDFSKNGSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKEAEKK-LPSD-- 970
      : | | | | | | | | : | | | : | : | :
Db      2039 DETPPTSVSESAPSQTDSV---PPETE-----ECPSITADANLDSEDESETIPTDKT 2088

Qy      971 -----TEKEDRSPS-----AIFSADLGKTSVVDDLLYWRDIKKTG 1004
      : | | | | | : | | | | | |
Db      2089 VTYKHMDPPPAPMQDRSPSPRHPDVSMVDPEALAI EQNLGKALKKDLKEKAKTKKPG 2145

```

# RESULT 15

## QRMSP1

microtubule-associated protein MAP1B - mouse

N;Alternate names: microtubule-associated protein MAP1(X); microtubule-associated protein MAP1.2; microtubule-associated protein MAP5

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 01-Sep-2000

C;Accession: S07549; S44387; A33645

R;Noble, M.; Lewis, S.A.; Cowan, N.J.

J. Cell Biol. 109, 3367-3376, 1989

A;Title: The microtubule binding domain of microtubule-associated protein MAP1B contains a repeated sequence motif unrelated to that of MAP2 and tau.

A;Reference number: A33645; MUID:90094539; PMID:2480963

A;Accession: S07549

A;Molecule type: mRNA

A;Residues: 1-2464 <NOB>

A;Cross-references: EMBL:X51396; NID:g52999; PIDN:CAA35761.1; PID:g53000

R;Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.

Arch. Biochem. Biophys. 310, 428-432, 1994

A;Title: Binding of heat-shock protein 70 (hsp70) to tubulin.

A;Reference number: S44387; MUID:94234720; PMID:8179328

A;Accession: S44387

A;Status: preliminary

A;Molecule type: protein

A;Residues: 653-663, 'IC' <SAN>

C;Superfamily: microtubule-associated protein MAP1B

C;Keywords: microtubule binding; phosphoprotein; tandem repeat

F;589-786/Domain: microtubule binding #status experimental <MTB>

F;589-592,639-642,649-652,655-658,660-663,668-671,674-677,679-682,683-686,687-690,691-694,695-698,699-702,708-711,712-715,716-719,720-723,727-730,758-761,764-767,783-786/Region: 4-residue repeats (K/R-K-E/D-X)

F;1861-2064/Region: 17-residue repeats

F;91,116,351,888,1124,1153,1168,1208,1662,1877,1918,2003,2030,2054,2083/Binding site: phosphate (Ser) (covalent) #status predicted  
 F;147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr) (covalent) #status predicted  
 F;1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 4.9%; Score 289; DB 1; Length 2464;  
 Best Local Similarity 19.8%; Pred. No. 0.0005;  
 Matches 262; Conservative 148; Mismatches 474; Indels 442; Gaps 55;

Qy	32	EDEEEEEEEEEDEDEDL-EELEVLERKPAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAP	90
		:        :       :	
Db	1009	EAEQSEEEGEEEDKAEDAREEGYEPDKTEADYVMAVADKAAEAGVTEEQYGY-----	1061
Qy	91	RGPLPAAPPV-APERQPSWDPSVSPSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPA	149
		: :     :     :	
Db	1062	LGTSAKQPGIQSPSREPA--SSIHDETLPGGSESEATA-----SDEENREDQPEEFTA	1112
Qy	150	SVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQED	209
		: :          :	
Db	1113	TSGYTQSTIEISSEPTPMDEMSTP-----RDVMSDETNNETESPSQEF	1156
Qy	210	FPSVLLETAASXPSLSPLSAASF-----EHEYLGNLSTVLPTEGTLQENVSE	257
		:   :      :     :     :	
Db	1157	VNITKYESSLYSQEYSKPAVASFNGLSEGSKTDATDGKDYNASASTISPPSSMEEDKFSK	1216
Qy	258	AS-----KEVSEK-----AKTLLIDRDLTEFS	279
		: :   :    :       :	
Db	1217	SALRDAYCSEEEKELKASAELDIKDVSDERLSPAKSPSLSPSPSPPIEKTPLGERSV-NFS	1275
Qy	280	---ELEYSEMGSFVSVPKAESA---VANPRE---EIIIVKNK-----	314
		:     :        :     : :	
Db	1276	LTPNEIKVSAEGEARSVSPGVTQAVVEEHCASPEEKTLEVVSQSVTGSAGHTPPYYQSP	1335
Qy	315	-DEEEKLVSNNILHXQQELPTALTCLKVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYADF	373
		: : :   :   :         :	
Db	1336	TDEKSSHLPTEVSENAQAVPVSF-----EFSEAKDE-NER--ASLSPMDEPVPDS	1382
Qy	374	K-PFERVWE-----VKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDSE	425
		:     : : :   :     :	
Db	1383	ESPVEKVLSPLRSPPLLGSESPYEDFLSADSKVLGR-----RSESPFE	1425
Qy	426	SSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPILLEDPTSENXT-----	478
		:   :   :     :	
Db	1426	GKNGKQGFDPRESPVSD-----LTSTGLYQDKQEEKSTGFIPIKEDFGPEKKTSDVETMS	1480
Qy	479	-----DEKKI-----EEKKAQIVTEKNTSTKTSNPFVAAQDSETD	514
		: : :   :     :     :	
Db	1481	SQSALALDERKLGGDVSPQTQIDVSQFGSFKEDTKMSISEGTVSDKSATP--VDEGVAEDT	1538
Qy	515	YVTTDNLTQVTEEVVA--NMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVM	572
		: :   :    :        : :   : : :	
Db	1539	YSHMEGVASVSTASVATSSFPPEPTDD-VSPSLHAEVGSHPSTEVDDSLSVSVVQTPTTF	1597
Qy	573	QES-LYPAAQLCP-----SFEESEATPSVLPDIVMEAPLNSAV--	610
		: :   :     : :   :   :	
Db	1598	QETEMSPSKEECPRPMSISPPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSFAMD	1657

Qy 611 -----PSAGASVIQ-----PSSSPLEASSVNYESIKHE--PENPPPY-----EEAM 649  
 | : |||| : | : | : | : | :  
 Db 1658 FSRQSPDHPTLGASVLHITENGPTVDYSPCDIQDSSLSHKIPPTTEPSYTDNDLSELI 1717

Qy 650 SVS-LKVSGIKEEIKEPENINAALQETE-----APYISIACDLIKE---TKLSAE 695  
 ||| : : | | : ||| : | : | : : : ||| :  
 Db 1718 SVSQVEASPSTSSAHTPSQIASPLQEDTLDVVPREMSLYASLASEKVQSLEGEKLSPK 1777

Qy 696 P-----APDFSDYSEMAK-----VEQPVPDHSELVEDS- 723  
 : | ||| : || : | | : : | :  
 Db 1778 SDISPLTPRESSPLYSPGFSDDSTSAAKETAAAHQASSSPPIDAATAEPYGRSSMLFDTM 1837

Qy 724 -----SPDSEPVDLFSD-----DS 737  
 ||| | | | |  
 Db 1838 QHHLALNRDLTTSSVEKDSGGKTPGDFNYAYQKPENAAAGSPDEEDYDYESQEKTIIRTHDV 1897

Qy 738 IPDVPQKQDETVML-----VKESLTETSFESMIEYENKEKLSALPPEGKPYLESF 788  
 : : | : | : | : | : | : | : | :  
 Db 1898 VRYYYEKTERTIKSPCDSGYSYETIEKTTKTPEDGGYTCEITEKTTTRTPEEGGYSY---- 1953

Qy 789 KLSLDNTKDTLLPDEVSS--TLSKKEKIPLQMEELSTAVYSNDD----- 829  
 : : : | ||| | | : : : : | :  
 Db 1954 ----EISEKTTRTPEVSGYTYEKTERSRRLDDISNGYDDTEDGGHTLGDCSYSYETTEK 2009

Qy 830 --LFISKEAQIRETETFSDDSP-----IEIIDEFPTLISSKTDSFSLAREYT----- 875  
 | | : ||| : || : : : | : | :  
 Db 2010 ITSFPESSESYSYETSTKTTRSPDTSAYCYETMEKITKTPQASTYSYETSDRCYTTEKKSP 2069

Qy 876 -----DL-----EVSH-KSEIANA---PDG---AGSLPCTELPHDLSLKNIQP--- 911  
 || | | : : : | : || | | : |  
 Db 2070 SEARQDVDLCLVSSCEFKHPKTELSPSFINPNPLEWFAGEEPTEESEKPLTQSGGAPPPS 2129

Qy 912 -----KVEEKISFSDDFSNGSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKEAE 964  
 : : | | | | | : | | : : | |  
 Db 2130 GKGQQRQCDDETPTTSVSESAPSQTDSDV---PPETE-----ECPSITADANIDSEDE 2179

Qy 965 KK-LPSD-----TEKEDRSPS-----AIFSADLGKTSVVDLLYWR 998  
 : : | : | : |||| : ||| |  
 Db 2180 SETIPTDKTVTYKHMDPPPAPMQDRSPSPRHPDVSMVDPDALAVDQNLGKAVKKDLKEKT 2239

Qy 999 DIKKTG 1004  
 || |  
 Db 2240 KTKKPG 2245

Search completed: January 22, 2004, 16:32:56  
 Job time : 45.1516 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:31:15 ; Search time 61.624 Seconds  
(without alignments)  
4932.919 Million cell updates/sec

Title: US-09-830-972-29  
Perfect score: 5923  
Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	% Query	Length	DB ID	Description
-----							

1	4851	81.9	986	4	Q8IUA4	Q8iua4 homo sapien
2	4289	72.4	1162	11	Q8BGM9	Q8bgm9 mus musculu
3	4277.5	72.2	1163	11	Q8K3G8	Q8k3g8 mus musculu
4	3658.5	61.8	1046	11	Q8BGK7	Q8bgk7 mus musculu
5	2527	42.7	639	11	Q8K290	Q8k290 mus musculu
6	1531	25.8	392	4	Q96B16	Q96b16 homo sapien
7	1201.5	20.3	375	11	Q8BHF5	Q8bhf5 mus musculu
8	1163	19.6	356	11	Q8BH78	Q8bh78 mus musculu
9	1152.5	19.5	357	11	Q8K3G7	Q8k3g7 mus musculu
10	867	14.6	179	6	Q9GM33	Q9gm33 macaca fasc
11	788.5	13.3	760	13	Q90638	Q90638 gallus gall
12	779	13.2	780	11	Q8K4S4	Q8k4s4 mus musculu
13	778	13.1	780	11	Q8K0T0	Q8k0t0 mus musculu
14	699	11.8	643	11	Q8CCU2	Q8ccu2 mus musculu
15	688	11.6	199	4	Q9BQ59	Q9bq59 homo sapien
16	672	11.3	208	13	Q90637	Q90637 gallus gall
17	671	11.3	267	11	Q63765	Q63765 rattus sp.
18	626	10.6	237	11	Q8C6D5	Q8c6d5 mus musculu
19	623.5	10.5	236	11	Q8VBU0	Q8vbu0 rattus norv
20	566	9.6	595	5	Q9VMV9	Q9vmv9 drosophila
21	528	8.9	234	5	Q9VMW3	Q9vmw3 drosophila
22	528	8.9	2484	5	Q9U347	Q9u347 caenorhabdi
23	526	8.9	224	5	Q9VMW1	Q9vmw1 drosophila
24	520	8.8	222	5	Q9VMW4	Q9vmw4 drosophila
25	519.5	8.8	2607	5	Q23187	Q23187 caenorhabdi
26	518	8.7	202	5	Q9VMW2	Q9vmw2 drosophila
27	357.5	6.0	2768	5	Q9VC00	Q9vc00 drosophila
28	341.5	5.8	5412	5	Q9W596	Q9w596 drosophila
29	340.5	5.7	222	5	Q23188	Q23188 caenorhabdi
30	339	5.7	1150	5	Q8IMM6	Q8imm6 drosophila
31	334	5.6	16215	5	Q9NFS3	Q9nfs3 drosophila
32	334	5.6	18074	5	Q9I7U4	Q9i7u4 drosophila
33	331	5.6	7962	4	Q10465	Q10465 homo sapien
34	331	5.6	34350	4	Q8WZ42	Q8wz42 homo sapien
35	325.5	5.5	5327	5	O76891	O76891 drosophila
36	322.5	5.4	846	5	Q8MRP6	Q8mrp6 drosophila
37	322.5	5.4	1109	5	Q9VAY4	Q9vay4 drosophila
38	318	5.4	864	5	Q9VGC8	Q9vgc8 drosophila
39	317.5	5.4	4969	11	Q8CF91	Q8cf91 mus musculu
40	317.5	5.4	5165	11	Q8CF92	Q8cf92 mus musculu
41	316	5.3	842	5	Q9VGC9	Q9vgc9 drosophila
42	315	5.3	10578	5	Q8ISF5	Q8isf5 caenorhabdi
43	315	5.3	18519	5	Q8ISF6	Q8isf6 caenorhabdi
44	315	5.3	18534	5	Q8ISF7	Q8isf7 caenorhabdi
45	314	5.3	864	5	Q95U45	Q95u45 drosophila

# ALIGNMENTS

## RESULT 1

Q8IUA4

ID Q8IUA4 PRELIMINARY; PRT; 986 AA.  
AC Q8IUA4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)



```

DE      RNT4 (RTN4).
GN      RTN4.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Oertle T., van der Putten H., Schwab M.E.;
RT      "Genomic Structure and Functional Characterization of the Promoter
RT      Structures of Human and Mouse Nogo/Rtn-4.";
RL      Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Oertle T., Schwab M.E.;
RL      Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Van der Putten H.;
RL      Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Testis;
RX      MEDLINE=22376540; PubMed=12488097;
RA      Oertle T., van der Putten H., Schwab M.E.;
RT      "Genomic Structure and Functional Characterization of the Promoter
RT      Structures of Human and Mouse Nogo/Rtn-4.";
RL      J. Mol. Biol. 325:299-323(2003).
DR      EMBL; AY102285; AAM64244.1; -.
DR      EMBL; AY123245; AAM64249.1; -.
DR      EMBL; AY123246; AAM64250.1; -.
DR      EMBL; AY123247; AAM64251.1; -.
DR      EMBL; AY123248; AAM64252.1; -.
DR      EMBL; AY123249; AAM64253.1; -.
DR      EMBL; AY123250; AAM64254.1; -.
SQ      SEQUENCE      986 AA;  108449 MW;  0CDE8F647036415A CRC64;

```

[illegible]

Qy	433	FPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDEKKIEEKKAQIVT	492
Db	241	FPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLGDPPTSENKTDEKKIEEKKAQIVT	300
Qy	493	EKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEV	552
Db	301	EKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEV	360
Qy	553	TGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPS	612
Db	361	TGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPS	420
Qy	613	AGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEAMSVSL-KVSGIKEEIKEPENINAA	671
Db	421	AGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKKVSGIKEEIKEPENINAA	480
Qy	672	LQETEAPYISIACDLIKETKLSAEPAPDFS DYSEMAKVEQPVPDHSSELVEDSSPDSEPVD	731
Db	481	LQETEAPYISIACDLIKETKLSAEPAPDFS DYSEMAKVEQPVPDHSSELVEDSSPDSEPVD	540
Qy	732	LFSDDSIPDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLS	791
Db	541	LFSDDSIPDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLS	600
Qy	792	LDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQIRETETFS DSSPIE	851
Db	601	LDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQIRETETFS DSSPIE	660
Qy	852	IIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQP	911
Db	661	IIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQP	720
Qy	912	KVEEKISFSDDFSKNGSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKEAEKKLPSDT	971
Db	721	KVEEKISFSDDFSKNGSATSKVLLLPDVSALA-TQAEIESIVKPKVLVKEAEKKLPSDT	779
Qy	972	EKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIAL	1031
Db	780	EKEDRSPSAIFS AELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIAL	839
Qy	1032	ALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKE	1091
Db	840	ALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKE	899
Qy	1092	LRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHY	1151
Db	900	LRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHY	959
Qy	1152	LGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	960	LGLANKNVKDAMAKIQAKIPGLKRKAE	986

RESULT 2

Q8BGM9

ID Q8BGM9 PRELIMINARY; PRT; 1162 AA.  
AC Q8BGM9;



Db 224 PLSTVSFKEHGYLGNL SAVASTE GTIEETLNEASRELPERATNPFVNRESAEFSVLEYSE 283

Qy 286 MGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEV 345  
 |||||: ||| |||:| | :||:| |:|: || : || || | |||:|||| |

Db 284 MGSSFNKSPKGESAMLVENTKEEVI VRSKDKED-LVCSAALHNPQESPATLTQVVKEDGV 342

Qy 346 VSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLES 404  
 :| || | ||| :|:| ||:||||||| |: ||||: | | |:||| :|:|

Db 343 MSPEKTMDFNEMKMSVVPVREEYADFKPFQAWEVKDTYEGSRDVLAA---RANMES 398

Qy 405 KVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATN 464  
 ||||| |||| | |||| |: || ||| :|| | |||| | :|:|||| |

Db 399 KVDKKCFEDSLEQKGHGKDSERNENASFPRTPELVKDGSRAYITCDSFS-SATESTAAN 457

Qy 465 IFPLLEDPTSENXTDEKKIEEKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKV 524  
 |||:||| |||| | |||| |: ||| || ||||| || ||||| |||||:|

Db 458 IFPVLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFVLVAIHDSEADYVTTDNLSKV 516

Qy 525 TEEVVANMPEGLTPDLVQEACESELNEVTGKTIAYETKMDLVQTSEVMQESLYPAAQLCP 584  
 || ||| ||||| ||||| ||||| ||||| |: ||||| :|||:| |||||

Db 517 TEAVVATMPEGLTPDLVQEACESELNEATGKTIAYETKVDLVQTSEAIQESIYPTAQLCP 576

Qy 585 SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPENPP 643  
 ||||:||||||| ||||| :|| |||| |||:|||| | |:|: || |||||

Db 577 SFEEAEATPSPVLPDIVMEAPLNSLLPSTGASVAQSPASPLEVPSVPVSYDGIKLEPENPP 636

Qy 644 PYEEAMSVSLKVGSIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDY 703  
 ||||| |:| | ||||| |: ||| || ||||| ||||| |||:|:| |

Db 637 PYEEAMSVALKTSDSKEEIKEPESFNAAQEAEPYISIACDLIKETKLSTEPSPEFSNY 696

Qy 704 SEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFS 763  
 |:| | |: |||| |:| |:| |:| ||||| |:| |:| |:|:|||| | |:

Db 697 SEIAKFEKSVDPDHCELVDDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVS-ET 755

Qy 764 MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTA 823  
 : :|:| |:| | | ||||| |:| ||| :|: |:| | |||| |:|

Db 756 VTQHKHKERLSASPQEVGKPYLESFQPNLHITKDA-ASNEIPTLTCKETISLOMEEFNTA 814

Qy 824 VYSNDDLFISKEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKS 883  
 :| |||| | ||| :|:| |:| ||||| ||||| |:| | | :| |||||:|

Db 815 IYSNDDLLSSKEDKMKESETFSDSSPIEIIDEFPTFVSAKDDS---PKEYTDLEVSNKS 870

Qy 884 EIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVLLLPDVSAL 943  
 |||| | |||:| ||| || | || : ||:| |:| |: ||| | |:|

Db 871 EIANVQSGANSLPCSELPCDLSFKNTYPKDEAHV--SDEFKSRSSVSKVPLLLPNVSAL 928

Qy 944 GHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKT 1003  
 :| |: :| |||| | |||:| ||||| |:| |:| | ||||| |||||

Db 929 -ESQIEMGNI VKPKVLTKEAEKKLPDTEKEDRSLTAVLSAELNKTSVVDLLYWRDIKKT 987

Qy 1004 GVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYL 1063  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 988 GVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYL 1047

Qy 1064 ESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLKFVFLMWVFTYVGALFNG 1123  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1048 ESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLKFVFLMWVFTYVGALFNG 1107

Qy 1124 LTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKI PGLKRKAE 1178  
| | | | | : | | | | | : | | | | |  
Db 1108 LTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQAKI PGLKRKAE 1162

### RESULT 3

O8K3G8

```

ID      Q8K3G8                PRELIMINARY;          PRT;   1163 AA.
AC      Q8K3G8;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Nogo-A.
GN      RTN4.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c;
RA      Jin W., Long M., Li R., Ju G.;
RT      "Cloning and expression of the mouse Nogo-A protein.";
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY114152; AAM77068.1; -.
DR      MGD; MGI:1915835; Rtn4.
DR      InterPro; IPR003388; Reticulon.
DR      Pfam; PF02453; Reticulon; 1.
DR      PROSITE; PS50845; RETICULON; 1.
SQ      SEQUENCE      1163 AA;  126691 MW;  6B5F362799417EA4 CRC64;

```

Query Match 72.2%; Score 4277.5; DB 11; Length 1163;  
Best Local Similarity 73.8%; Pred. No. 3.9e-205;  
Matches 883; Conservative 108; Mismatches 154; Indels 51; Gaps 19;

Qy	1	MEDLDQSPVLVSS--DSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKP	59
		:                           : : : : : : : : : : : : :	
Db	1	MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEDEDEDEDEDEDLEELEVLERKP	60
Qy	60	AAGLSAAPVPTAPAAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSVSTVPA	119
		:    :                         :     :	
Db	61	AAGLSAVPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAAS---A	114
Qy	120	PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRR	179
Db	115	PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAE-----PAAPPSTPAAPKRR	163
Qy	180	GSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXP SLS	225
		:                  : : : : : : : : : : : : :	
Db	164	GSGSVDET L FALPAASEPVI PSSAEKIMDLKEQPGNTVSSGQEDFPSVL FETAASLPSLS	223
Qy	226	PLSAASFKEHEYLG N LSTVLPT EGT LQENVSEASKEVSEKAKTLLID RDLTEFSELEYSE	285
		:  : : : : : : : : : : : :	
Db	224	PLSTVSFKEHGYLG NLSAVASTEGTIEETLNEASRELPERATNPFVNRESAEFSVLEYSE	283
Qy	286	MGSSFSVSPKAESAVI VANPREEII VKNKDEEEKLVSN N I LHXQOELPTALT KLVKEDEV	345

Db	284	MGSSFNNGSPKGESAMLVENTKEEVI VRSKDKED-LVCSAALHNPQESPATLTQVVKEDGV	342
Qy	346	VSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLES	404
Db	343	MSPEKTMDFNEMKMSVVPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANMES	398
Qy	405	KVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATN	464
Db	399	KVDKKCFEDSLEQKSHGKDESERNENASFPSTPELVKDGSRAYITCDSFT-SATESTAN	457
Qy	465	IFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTQV	524
Db	458	IFPVLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAIHDSEADYVTTDNLSKV	516
Qy	525	TEEVVANMPEGLTPDLVQEACESELNEVTGKIAYETKMDLVQTSEVMQESLYPAAQLCP	584
Db	517	TEAVVATMPEGLTPDLVQEACESELNEATGKIAYETKVDLVQTSEAIQESIYPTAQLCP	576
Qy	585	SFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPENPP	643
Db	577	SFEEAEATPSPVLPDIVMEAPLNSLLPSTGASVAQPSASPLEVSPSPSYDGIKLEPENPP	636
Qy	644	PYEEAMSVSLKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDY	703
Db	637	PYEEAMSVALKTSDAKEEIKEPESFNAAQEAAPYISIACDLIKETKLSTEPSPGFSNY	696
Qy	704	SEMAKVEQPVPDHSELVEDSSPDSEPVDFLSDDSIPDVPQKQDETVMVLKESLTETSFES	763
Db	697	SEIAKFEKSVDPHCELVDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVS-ET	755
Qy	764	MIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTA	823
Db	756	VTQHKHKERLSASPQEVGKPYLESFQPNLHITKDA-ASNEIPTLTKKETISLQMEEFNTA	814
Qy	824	VYSNDDLFIKSEAQIRETETFSDDSPIEIIDEFPTLISSKTDSEFSLAREYTDLEVSHKS	883
Db	815	IYSNDDLSSKEDKMKESETFSDSSPIEIIDEFPTFVSAKDDS----PKEYTDLEVSNKS	870
Qy	884	EIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVLLLPPDVSAL	943
Db	871	EIANVQSGANSLPCSELPCDLSFKNTYPKDEAHV--SDEFKSRSSVSKVPLLLPNVSAL	928
Qy	944	GHTQAEIESIVKPKVLEKEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKT	1003
Db	929	-ESQIEMGNIVKPKVLTKEAEKKLPSDTEKEDRSLTAVLSAELNKTSVVDLLYWRDIKKT	987
Qy	1004	GVV-FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1062
Db	988	GVVYFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1047
Qy	1063	LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFN	1122
Db	1048	LESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFN	1107
Qy	1123	GLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIAKIPGLKRKAE	1178

Db 1108 GLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 1163

RESULT 4

Q8BGK7

ID Q8BGK7 PRELIMINARY; PRT; 1046 AA.  
AC Q8BGK7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE RTN4.  
GN RTN4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7, and 129SvcJ7;  
RA Oertle T., van der Putten H., Schwab M.E.;  
RT "Genomic Structure and Functional Characterization of the Promoter  
RT Structures of Human and Mouse Nogo/Rtn-4.";  
RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7, and 129SvcJ7;  
RA Oertle T., Schwab M.E.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7;  
RA Van der Putten H.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SvcJ7;  
RA Van der Putten H., Mir A.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY102280; AAM73502.1; -.  
DR EMBL; AY102286; AAM73507.1; -.  
SQ SEQUENCE 1046 AA; 114221 MW; 8CE2E2238ED51222 CRC64;

Query Match 61.8%; Score 3658.5; DB 11; Length 1046;  
Best Local Similarity 69.5%; Pred. No. 2.6e-174;  
Matches 767; Conservative 108; Mismatches 148; Indels 81; Gaps 17;

Qy 97 APPVAPERQ-----PSWDPS-----PVSSTVPAP---SPLSAAVSPSKLPED 136  
|||:| | :| || | || | |||  
Db 2 APPLAGGGQKGAASEAWVPSLFGVSGSTCTAAKSLVPIPARSSRLSAA----- 51  
  
Qy 137 DEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKRRGSSGAVVXXXXKIMDLK 196  
: | :: || | : | | :: : |||||  
Db 52 -----RNETLFALPA---ASEFVIPSSAE-----KIMDLK 78  
  
Qy 197 EQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAAFKEHEYLGNLSTVLPTEGTLQENVS 256  
|||||:|:||||||| ||||| ||||| |||||:| ::  
Db 79 EQPGNTVSSGQEDFPSVLFETAASLPSLSPLSTVSFKEHGYLGNLSAVASTEGTIEETLN 138

Qy 257 EASKEVSEKAKTLLIDRDLTEFSELEYSEMSSFSVSPKAESAVIVANPREEIIIVKNKDE 316  
 |||:|: |:| :::| ||| |||||: ||| |||:| |:|:|:|:|:  
 Db 139 EASRELPERATNPFVNRESAEFSVLEYSEMSSFNKSPKGESAMLVENTKEEVIVRSKDK 198

Qy 317 EEKLVSNNILHXQQELPTALTCLVKEDEVSSSEKAKDSFNEKRVAVEAPMREEYADFKPF 376  
 |: || : || | |||:| ||| |:| || | ||| :::| |:| ||||| ||||  
 Db 199 ED-LVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDIFNEMKMSVVAPVREEYADFKPF 257

Qy 377 ERVWEVKDSKEDS-DMLAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSSESSNDDTSFPS 435  
 |: |||||: | | |:||| :|:| ||||| |||| | |||| |:| |||  
 Db 258 EQAWEVKDITYEGSRDLAA----RANMESKVDDKKCFEDSLEQKGHGKDSSESRNENASFP 313

Qy 436 TPEGIKDRSGAYITCAPFNPAATESIATNIFPLEDPTSENXTDEKKIEEKKAQIVTEKN 495  
 ||| :|| | |||| |: :||| | ||||| |||| | |||||:| |||:| |||  
 Db 314 TPELVKDGSRAYITCDSFS-SATESTAANIFPVLEDHTSENKTDEKKIEERKAQIITEK- 371

Qy 496 TSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDLVQEACESELNEVTGT 555  
 || ||||| || ||| |||||: ||| ||| ||||| ||||| ||||| ||||| |||||  
 Db 372 TSPKTSNPFVVAIHDSADYVTTDNLTKVTEAVVATMPEGLTPDLVQEACESELNEATGT 431

Qy 556 KIIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGA 615  
 |||||: ||||| :|||: || |||||: ||||| ||||| ||||| :|| ||  
 Db 432 KIIAYETKMDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSTGA 491

Qy 616 SVIQPSSSPLEA-SSVNYESIKHEPENPPPYEEAMSVSLKVSGIKEEIKEPENINAAALQE 674  
 || |||: ||| | |:| || ||||| |||||: || | |||||: ||| ||  
 Db 492 SVAQPSASPLEVPSVSYDGIKLEPENPPPYEEAMSVALKTSDSKEEIKEPESFNAAAQE 551

Qy 675 TEAPYISIIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFS 734  
 ||||| ||||| ||||| |||:|:|:|:| ||| |: ||| |||:|:|:| |||||  
 Db 552 AEAPYISIIACDLIKETKLSPEFSNYSEIAKFEKSVDPHCELVDSSPESEPVDLFS 611

Qy 735 DDSIPDVPQKQDETVMVLKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDN 794  
 |||||: ||| |:| |||: ||||| | |:| :::|:|:| ||| | |||||: |:|  
 Db 612 DDSIPEVPQTQEEAVMLMKESLTVS-ETVTQHKHKERLSASPQEVGKPYLESFQPNLHI 670

Qy 795 TKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFIKSEAQIRETETFSOSSPIEIID 854  
 ||| :|: |:|:| ||| ||||| :|:| ||||| ||| :::|:| ||||| |||||  
 Db 671 TKDA-ASNEIPTLTKKETISLQMEEFNTAISNDDLSSKEDKMKESSETFSOSSPIEIID 729

Qy 855 EFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVE 914  
 |||| |:| || |:| |||||: ||||| || |||||: ||| || ||| |||  
 Db 730 EFPTFVSAKDDS----PKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNTYPKDE 785

Qy 915 EKISFSDDFSKNGSATSQVLLLPDVSALGHTQAEIESIVKPKVLEKEAEKKLPDTEKE 974  
 : ||:|: |:| ||| || |:| ||| :| |:|:| ||||| ||||| ||||| |||||  
 Db 786 AHV--SDEFSKSRSSSVKVPPLLPNVSALESGIEMGNIVKPKVLTKEAEKKLPDTEKE 842

Qy 975 DRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALL 1034  
 ||| :|: |:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 843 DRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALL 902

Qy 1035 SVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRR 1094  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 903 SVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRR 962



```
Qy      1095 LFLVDDLVDSLKFAVLMMWVFYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGL 1154
          |||||||
Db       963 LFLVDDLVDSLKFAVLMMWVFYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGL 1022
          |||||
Qy      1155 ANKNVKDAMAKIQAKIPGLKRKAEE 1178
          |||:|||||
Db     1023 ANKSVKDAMAKIQAKIPGLKRKAEE 1046
```

## RESULT 5

```
Q8K290
ID   Q8K290          PRELIMINARY;          PRT;   639 AA.
AC   Q8K290;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC032192; AAH32192.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Hypothetical protein.
SQ   SEQUENCE      639 AA;  70312 MW;  309A19DA37603F11 CRC64;
```

Query Match 42.7%; Score 2527; DB 11; Length 639;  
Best Local Similarity 79.5%; Pred. No. 3.6e-118;  
Matches 515; Conservative 57; Mismatches 66; Indels 10; Gaps 6;

Qy	532	MPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEA	591
Db	1	MPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEA	60
Qy	592	TPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESI KHEPENPPPPYEEAMS	650
Db	61	TPSPVLPDIVMEAPLNSLLPSTGASVAQPSASPLEVPSPVSYDGIKLEPENPPPPYEEAMS	120
Qy	651	VSLKVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSEMAKVE	710
Db	121	VALKTSDAKEEIKEPESFNAAQEAAPYISIIACDLIKETKLSTEPSPEFSNYSEIAKFE	180
Qy	711	QPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVM LVKESLTETSFESMIEYENK	770
		:	
Db	181	KSPVDHCELVDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVS-ETVTQHKHK	239
Qy	771	EKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDL	830
		:	
Db	240	ERLSASPQEVGKPYLESFPQNLHITKDA-ASNEIPTLTKKETISLQMEEFNTAIYSNDDL	298

Qy 831 FISKEAQIRETETFSDDSSPIEIIDFPTLISSKTD SFSKLAREYTDLEVSHKSEIANAPD 890  
 ||| ::|:||||||||||||||| :|:| || :|||||||:|||||||  
 Db 299 LSSKEDKMKESETFSDSSPIEIIDFPTFVSAKDDS----PKEYTDLEVS NKSEIANVQS 354

Qy 891 GAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SALGHTQAEI 950  
 || ||||:||| ||| || || | : ||:||||: |: ||| || |:|||| :| ||:  
 Db 355 GANSLPCSELP CDLSFKNTYPKDEAHV--SDEF SKSRSSVSKVPLLLPNVSAL-ESQIEM 411

Qy 951 ESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGVVFGAS 1010  
 :||||||| ||||:||||||||||| :|: ||:| |||||||||||||  
 Db 412 GNIVKPKVLTKEAEKKLP SDTEKEDRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGAS 471

Qy 1011 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS 1070  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 472 LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS 531

Qy 1071 EELVQKYSNSALGHVNCTIKELRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILA 1130  
 ||||||||||||||||| |||||||||||||||||||||||||||||||||  
 Db 532 EELVQKYSNSALGHVNSTIKELRRFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILA 591

Qy 1131 LISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 |||||:|||||||||||||||:|||||||||||||||  
 Db 592 LISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 639

RESULT 6

Q96B16

ID Q96B16 PRELIMINARY; PRT; 392 AA.  
 AC Q96B16;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein (RTN4).  
 GN RTN4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Oertle T., van der Putten H., Schwab M.E.;  
 RT "Genomic Structure and Functional Characterization of the Promoter  
 RT Structures of Human and Mouse Nogo/Rtn-4.";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Oertle T., Schwab M.E.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Van der Putten H.;









ID Q8BH78 PRELIMINARY; PRT; 356 AA.  
AC Q8BH78;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE RTN4.  
GN RTN4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7, and 129SvcJ7;  
RA Oertle T., van der Putten H., Schwab M.E.;  
RT "Genomic Structure and Functional Characterization of the Promoter  
RT Structures of Human and Mouse Nogo/Rtn-4.";  
RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7, and 129SvcJ7;  
RA Oertle T., Schwab M.E.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvcJ7;  
RA Van der Putten H.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SvcJ7;  
RA Van der Putten H., Mir A.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY102281; AAM73503.1; -.  
DR EMBL; AY102286; AAM73508.1; -.  
SQ SEQUENCE 356 AA; 38403 MW; 4366C03BA9630B56 CRC64;

Query Match 19.6%; Score 1163; DB 11; Length 356;  
Best Local Similarity 27.7%; Pred. No. 1.8e-50;  
Matches 326; Conservative 12; Mismatches 17; Indels 824; Gaps 8;

Qy 1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKP 59  
|||:|||| ||||| ||||| ||||| |||||:||:||||:|||||||  
Db 1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEDEDEDEDEDLEELEVLERKP 60  
Qy 60 AAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVPA 119  
||||||| ||| |||:| |:| ||||| ||||| |||||: || :| |  
Db 61 AAGLSAAPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAAS---A 114  
Qy 120 PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPSTPAAPKRR 179  
|| ||||| ||||| ||||| || || || || ||||| |||||  
Db 115 PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAE-----PAAPPSTPAAPKRR 163  
Qy 180 GSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAASFKEHEYLG 239  
||  
Db 164 GS----- 165

Qy	240	NLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVSPKAESA	299
Db	166	-----	165
Qy	300	VIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEVSSSEKAKDSFNEKR	359
Db	166	-----	165
Qy	360	VAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTN	419
Db	166	-----	165
Qy	420	HEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTD	479
Db	166	-----	165
Qy	480	EKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTQVTEEVVANMP EGLTPD	539
Db	166	-----	165
Qy	540	LVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPD	599
Db	166	-----	165
Qy	600	IVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIK	659
Db	166	-----	165
Qy	660	EEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQPVPDHSEL	719
Db	166	-----	165
Qy	720	VEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMIEYENKEKLSALPPE	779
Db	166	-----	165
Qy	780	GGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFISKEAQIR	839
Db	166	-----	165
Qy	840	ETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTE	899
Db	166	-----  :-----GSV-----	168
Qy	900	LPHDLSLKNIQPKVEEKISFSDDFSKNGSATSQVLLLPDVSALGHTQAEIESIVKPKVL	959
Db	169	-----	168
Qy	960	EKEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTV	1019
Db	169	-----     -----VVDLLYWRDIKKTGVVFGASLFLLLSLTV	197
Qy	1020	FSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSN	1079
Db	198	FSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSN	257
Qy	1080	SALGHVNCTIKELRRLFLVDDLVDLSKFAVLMWVFTYVGALFNGLTLLIILALISLFSVPV	1139



Db	258	SALGHVNSTIKELRRFLVDDLVDLSKFVLMWVFTYVGALFNGLTLLILALISLFSIPV	317
Qy	1140	IYERHOAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	318	IYERHOAQIDHYLGLANKSVKDAMAKIOAKIPGLKRKAE	356

## RESULT 9

```
Q8K3G7
ID   Q8K3G7          PRELIMINARY;          PRT;   357 AA.
AC   Q8K3G7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Nogo-B.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Jin W., Li R., Long M., Shen J., Ju G.;
RT   "Cloning and expression of the mouse Nogo-B protein.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY114153; AAM77069.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE      357 AA;  38566 MW;  73BB3D17DFDBDF15 CRC64;
```

Query Match 19.5%; Score 1152.5; DB 11; Length 357;  
Best Local Similarity 27.6%; Pred. No. 6e-50;  
Matches 326; Conservative 12; Mismatches 17; Indels 825; Gaps 9;

Qy	1	MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKP	59
Db	1	MEDIDQSSLVSSSADSPRPAPAFKYQFVTEPEDEEDEDEDEDEDEDEDELEELEVLERKP	60
Qy	60	AAGLSAAPVPTAPAAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVPA	119
Db	61	AAGLSAAPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPPTAPERQPSWERSPAAS--A	114
Qy	120	PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRR	179
Db	115	PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAE-----PAAPPSTPAAPKRR	163
Qy	180	GSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAASFKEHEYLG	239
Db	164	GS-----	165
Qy	240	NLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFSVSPKAESA	299
Db	166	-----	165

Qy	300	VIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTKLVKEDEVVSSEKAKDSFNEKR	359
Db	166	-----	165
Qy	360	VAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTN	419
Db	166	-----	165
Qy	420	HEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPILLEDPTSENXTD	479
Db	166	-----	165
Qy	480	EKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPD	539
Db	166	-----	165
Qy	540	LVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPD	599
Db	166	-----	165
Qy	600	IVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIK	659
Db	166	-----	165
Qy	660	EEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSEL	719
Db	166	-----	165
Qy	720	VEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMIEYENKEKLSALPPE	779
Db	166	-----	165
Qy	780	GGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQIR	839
Db	166	-----	165
Qy	840	ETETFSDSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTE	899
Db	166	-----  : -GSV-----	168
Qy	900	LPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SALGHTQAEIESIVKPKVL	959
Db	169	-----	168
Qy	960	EKEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVV-FGASLFLLLSLT	1018
Db	169	-----            VVDLLYWRDIKKTGVVYFGASLFLLLSLT	197
Qy	1019	VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS	1078
Db	198	VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS	257
Qy	1079	NSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVP	1138
Db	258	NSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIP	317

Qy 1139 VIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 ||||||||||||||||:||||||||||||||  
 Db 318 VIYERHQAQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 357

RESULT 10

Q9GM33

ID Q9GM33 PRELIMINARY; PRT; 179 AA.  
 AC Q9GM33;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical 19.9 kDa protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AB049853; BAB16739.1; -.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 179 AA; 19949 MW; 5F8CD4383FEE9E02 CRC64;

Query Match 14.6%; Score 867; DB 6; Length 179;  
 Best Local Similarity 98.3%; Pred. No. 4.1e-36;  
 Matches 176; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1000 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHFP 1059  
 :|||||||||||||||||||||||||||||||||||||||||:||||||||||||||  
 Db 1 MKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYRGVIAIQKSDEGHFP 60  
 Qy 1060 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGA 1119  
 ||||||||||||||||||||||||||||||||||||||||:||||||||||||||  
 Db 61 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLRFVLMWVFTYVGA 120  
 Qy 1120 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 ||||||||||||||||||||||||||||||||||||||||:||||||||||||||  
 Db 121 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 179

RESULT 11

Q90638

ID Q90638 PRELIMINARY; PRT; 760 AA.  
 AC Q90638;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE ChS-Rex-b.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96386034; PubMed=8793864;  
 RA Baka I.D., Ninkina N.N., Pinon L.G., Adu J., Davies A.M.,  
 RA Georgiev G.P., Buchman V.L.;  
 RT "Intracellular compartmentalization of two differentially spliced s-  
 RT rex/NSP mRNAs in neurons."  
 RL Mol. Cell. Neurosci. 7:289-303(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97183663;  
 RA Ninkina N.N., Baka I.D., Buchman V.L.;  
 RT "Rat and chicken s-rex/NSP mRNA: nucleotide sequence of main  
 RT transcripts and expression of splice variants in rat tissues."  
 RL Gene 184:205-210(1997).  
 DR EMBL; U17606; AAC60075.1; -.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 SQ SEQUENCE 760 AA; 82502 MW; 465C1B429F799D5C CRC64;

Query Match 13.3%; Score 788.5; DB 13; Length 760;  
 Best Local Similarity 27.7%; Pred. No. 2.4e-31;  
 Matches 296; Conservative 115; Mismatches 315; Indels 341; Gaps 38;

Qy	143	PPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNT	202
		:	
Db	4	PPDPQDLLLAGTAER-W-----AAAGADEYAAGAALRDGDGA-----QQREQL	45
Qy	203	ISAGQEDFPSVLLETAASXPSLSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEV	262
		:     :           : :   :   : :	
Db	46	AFGSAREHPPVAMATA-----SPGVTASSRLFDY-----GSSSANGADSSFYT	88
Qy	263	SEKAKTLLID-----RDLTEFSELEYSEMGSFSVSPKAESAVIVANPREEIIVKNKDE	316
		:         : :   :     : :   : :	
Db	89	S-----LISDVHYTTPRDNTYFTGV-YQQENSPIPCSGSTEGFNALGHPVQDV---TGFE	139
Qy	317	EEKLVSNNILHXQQELPTALTKLVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFKPF	376
		: :       : :     :	
Db	140	SRGLFS-----LDSGIEMTPAES-----EVDKSLTDPMKVEGYKYMDI	178
Qy	377	ERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSESSNDD--TSFP	434
		:     :       :   :	
Db	179	SRPEEIK-----YQEKHDPDSEDESPDLIDEYR	206
Qy	435	STPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDEKKIEEKKAQIVTEK	494
		: :   : :   :   : :   :	
Db	207	GTPIG----SGH--AAEPQRTTASEAIKA---PKEQDPLE----DKSFRDQHNASVVTAP	253

Qy 495 NTSTKTSNPFFVAAQDSETDYVTTDNLTKVTEEVVANM----PE-----GLTP 538  
 Db 254 VKITLTETPGAREATSKEASVTQPKSGLKPSHEVVPTVMVSEPEDDSPGSVTPPSSGTEP 313

Qy 539 DLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLP 598  
 Db 314 SGSESQKGKSLSEDELISAIKEAKGFSFETSEVQQS---PAV----SAEKQEOKMKPGRP 366

Qy 599 DIVMEAPLNSAVPSA--GASVIQ-PSSSPLEASSV--NYESIKHEPENPPPYEEAMSVS 652  
 Db 367 --AVPSPLDNEASSAESGDSEIELVSEDPLAAEEVLHSNYMTFSH-IGGPPP--SPASPS 421

Qy 653 LKVSIGIKKEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQP 712  
 Db 422 IQYSILREERE-----AELDSELIIESCD---ASSASEESP-----KREQD 459

Qy 713 VPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMIEYENKEK 772  
 Db 460 SPLMKPMVMD-----IIKE-----ENSSR 478

Qy 773 LSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVYSNDDLFI 832  
 Db 479 ASADYE----- 485

Qy 833 SKEAQIRETETFS DSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGA 892  
 Db 486 -----ASKTTE-SRMNRE-----NLADSA 503

Qy 893 GSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDV SALGHTQAEIES 952  
 Db 504 SYLKSS-----FVAPKVSSE-----PPTSAVSTEELKERI 533

Qy 953 IVKPKVLEKEAEKKLPDTEKEDRSPSAI-FSADLGKTSVVDLLYWRDIKKTGVVFGASL 1011  
 Db 534 ILKKPIEETVNVQSKVSSKDSGKRSPLALPLL PFLNKQKAINLLYWRDIKQTGIVFGSLL 593

Qy 1012 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISE 1071  
 Db 594 LLLFSLTQFSVSVVAYLALAGLSATISFRIYKSVLQAVQKTDEGHPFKAYLDMEMNLSQ 653

Qy 1072 ELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILAL 1131  
 Db 654 DQIQKYTDCLQLYVNSTVKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLIMAV 713

Qy 1132 ISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 Db 714 VSMFTLPVVYDKYQAQIDQYLGLVRTHINTVVAKIQAKIPGAKRKAE 760

RESULT 12

Q8K4S4

ID Q8K4S4 PRELIMINARY; PRT; 780 AA.

AC Q8K4S4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Reticulon 1A.

GN RTN1 OR RTN-1A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR; TISSUE=Brain;  
 RA Hirata T., Nomura T., Takagi Y., Sato Y., Tomioka N., Fujisawa H.,  
 RA Osumi N.;  
 RT "Mosaic development of the olfactory cortex with Pax6-dependent and  
 RT -independent components.";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AB074899; BAB96551.1; -.  
 DR MGD; MGI:1933947; Rtn1.  
 DR InterPro; IPR001951; Histone\_H4.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS00047; HISTONE\_H4; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 SQ SEQUENCE 780 AA; 83504 MW; 545F5638C576A069 CRC64;

Query Match 13.2%; Score 779; DB 11; Length 780;  
 Best Local Similarity 30.8%; Pred. No. 7.5e-31;  
 Matches 236; Conservative 108; Mismatches 227; Indels 194; Gaps 28;

QY	538	PDLVQEACESELNEVTG-----TKIAYETKMDLVQTSEVMQ-ESLYPAAQLCPSFEES	589
		: :     :   : :   :   : :   :   :	
Db	86	PDALDHSPSSTLKDGEACYSLSIDVCYPPREDSSAYFTGILQKENGHITTSESP--EEP	143
QY	590	EATPSPVLPDIVMEAP---LNS-----AVPSAGASVIQSSSPLE---ASSVNY-----	632
		: :   :   :   :   :   :   :	
Db	144	E-TPGPSLPEVPGMEPQGLLSSDSGIEMTPAESTEVNKKILADPLDQMKAEAYKYIDITRP	202
QY	633	ESIKHEPENPPPYEE-----AMSVSLKVSIGIKKEIKEPENINAAL-----QETEA	677
		:   :     :       : : :   :   :	
Db	203	QEAKGQEEQHFGLEDKDLDFKDKGTEVSTKAEGVRAP-NQPAPVEGKLIKDHLEESTFA	261
QY	678	PYISIACDLIKETKLSAEPAPDFSDYSE----MAKVEQVPDPHSELVEDSSPDSEPVDL	732
		: : :     :   : : :   :   :	
Db	262	PYIDELSD--EQHRVSLVTAPVKITLTEIEPPLMTATQETIPEKQDLCLKPSPDTVPTVT	319
QY	733	FS---DDS-----IPDVPQKQ-----DETVMLVKESLTETSFESMIEYENKE	771
		:     : : :   :   :	
Db	320	VSEPEDDSPGSVTPPSSGTEPSAAESQGKGSVSEDELIAAIKEA-----KGLSYETTE	372
QY	772	KLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDL	831
		:     :     :     : : :	
Db	373	SPRPVGQVADKP-----KTKTRSGLPITPSPLDQEASSAESGD---	410
QY	832	ISKEAQIRETETFTSDSSPIEIIIDEFPTL-----ISSKTDSFSKLAREYTDL	877
		: :     :   : : :	
Db	411	-----SEIELVSEDPMASEDALPSGYVSFGHVSGPPPPSPASPSIQYSIL	454
QY	878	-----EVSHKSEIANAPDGAGSLP-----CTELPHDLS	905
		:	
Db	455	REEREAELDSELIIESCDASSASEESPKEQDSPMPKPGALDAIREETGSRATEERAPSH	514

Qy	906	LKNIQPKVEEKISFSDDFSKNGSATSKVLLLLPPDVSALGHTQAEIESIVKPKVLEKEAEK	965
Db	515	QGPVEP--DPMLSF-----APAAALQSRPEPSS----GDGASVPEP---PRSQQQ	555
Qy	966	KLPSDTEKEDRSPSAI-FSADLG-----KTSVVDLLYWRDIKKTGVVFGASLFL	1013
Db	556	KPEEEAVSSSQSPTATEIPGPLGSGLMPLPFFNKQKAIDLLYWRDIKQTGIVFGSFLLL	615
Qy	1014	LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEEL	1073
Db	616	LFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQ	675
Qy	1074	VQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALIS	1133
Db	676	IQKYTDCLQLYVNSTLKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALFNGLTLLLMMAVVS	735
Qy	1134	LFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	736	MFTLPVVVYVKHQAQVDQYLGVLVRTHINTVVAKIQAKIPGAKRHA	780

Q8K0T0

Query Match 13.1%; Score 778; DB 11; Length 780;  
Best Local Similarity 30.8%; Pred. No. 8.4e-31;  
Matches 236; Conservative 108; Mismatches 227; Indels 194; Gaps 28;





OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK032109; BAC27708.1; -.  
 SQ SEQUENCE 643 AA; 68769 MW; 15E3335B280533F8 CRC64;

Query Match 11.8%; Score 699; DB 11; Length 643;  
 Best Local Similarity 31.9%; Pred. No. 5.6e-27;  
 Matches 201; Conservative 95; Mismatches 186; Indels 148; Gaps 20;

QY	591	ATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMS	650
		:       : :     :     :     :	
Db	120	ATVEVTLPN-----LRGAWPNSVMGEVTEVDSSGESDDTVIEDITEKPDSLP-----S	167
QY	651	VSLKVGSIKEEIKPENINAALQETEAPYISIIACDLIKETKLSAEPAPDFSDYSEMAKVE	710
		:     :       : :     :     :     :	
Db	168	AAAKTS--EREIKETPS-----RETV-----	186
QY	711	QPVPDHSSELVEDS-SPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMIEYEN	769
		:     :   :     :     :     :     :	
Db	187	-----RSEMCENSEQPQAQP-----ETPTQKSLEGEVASQVPNTLNEVTPEKLDMTNN	234
QY	770	KEKLSALPP----EGG-----KPYLESFKLSLDNTKDTLLPDEVSTL--SKKEKIPLQ	816
		:                      :     :   :     :	
Db	235	PKVCSAAPPVSLNETGFSLTVPASAKLESLLGKYVEDTDGSSPEDLMAVLGTGAEEKGIVD	294
QY	817	MEELSTAVYSNDDLFIKSEAQIRETETFSDDSPIEII DEFPTLISSKT---DSFSKLAR	872
		: :   : :   :   :     :   :	
Db	295	KEE-----GDVL-----EAVLEKIADFKNTLPVELLHE-SELGSETKNIKSKYSEDSR	342
QY	873	EYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI-----	917
		:   :   :   :   :   :	
Db	343	ETT-----GGAPTMSPDLEQEQLT-IRA IKELGERQAEKVQDEGISSGGKLGKQTF	391
QY	918	-----SFSDDFSKNGSATSKVLLLLPPDV SALGHTQAE-IESIVKPKVLEKEAEKKLPS	969
		:   : : : :   :   :   :	
Db	392	APQSGPQSSSDILEHTDVKTGSDLGIPKNPTIIKNTRIDSISLTKTEMVNKNVLRLLS	451
QY	970	DTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGVVFGASLFLLSLTVFSIVSVTAYI	1029
		:   :   :   :   :   :   :   :	
Db	452	D-----FPVHDLIFWRDVKKTG FVFGTTLIMLLSLAAF SVISVVSYL	493
QY	1030	ALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTI	1089
		:          :       :       : : :     :   :	
Db	494	ILALLSVTISFRVYKSVIQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHVNKAL	553
QY	1090	KELRRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQID	1149

```

      | : |||||:||||| || ||: |||||:||||:||||| : :|||:|:|:|: |||
Db      554 KLIIRLFLVEDLVDSLKLAVFMWMLTYVGAVFNGITLLILAELLVFSVPVYKYKTQID 613

Qy      1150 HYLGLANKNVKDAMAKIQAKIPGL-KRKAE 1178
      ||:|:| | : |||||:|:|: |:|
Db      614 HYVGIARDQTKSIVEKIQAQLPGIAKKKAE 643

```

RESULT 15

Q9BQ59

```

ID   Q9BQ59          PRELIMINARY;          PRT;          199 AA.
AC   Q9BQ59;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to reticulon 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC003003; AAH03003.1; -.
DR   EMBL; BC000314; AAH00314.1; -.
DR   InterPro; IPR001951; Histone_H4.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS00047; HISTONE H4; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE 199 AA; 22642 MW; 7CFA44CC568DF6D8 CRC64;

```

```

Query Match          11.6%; Score 688; DB 4; Length 199;
Best Local Similarity 67.5%; Pred. No. 4e-27;
Matches 129; Conservative 32; Mismatches 30; Indels 0; Gaps 0;

```

```

Qy      988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKQVI 1047
      |: :|||:|:|:|:|:| | || ||| ||:| | |:| | | |||:|:|
Db      9 KSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVSVVAYLALAALSATISFRIYKSVL 68

Qy      1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKF 1107
      |:|:|:|:|:|:|:|:| | : :|:| :|:|:| :| | |:|:|:|:| | |||:|:|
Db      69 QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKF 128

Qy      1108 AVLMMWVF'TYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167
      |||||: |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      129 AVLMWLLTYVGALFNGLTLLLMAVVSMTFLPVVYVKHQAQIDQYLGVRTHINAVVAKIQ 188

Qy      1168 AKIPGLKRKAE 1178
      ||||| || |
Db      189 AKIPGAKRHAE 199

```

Search completed: January 22, 2004, 16:34:12  
Job time : 68.624 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:34:19 ; Search time 58.0516 Seconds  
(without alignments)  
4195.163 Million cell updates/sec

Title: US-09-830-972-29  
Perfect score: 5923  
Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description

1	5815	98.2	1192	9	US-09-758-140-6	Sequence 6, Appli
2	5815	98.2	1192	9	US-09-972-599A-6	Sequence 6, Appli
3	5815	98.2	1192	15	US-10-060-036-71	Sequence 71, Appl
4	5810	98.1	1192	9	US-09-789-386-2	Sequence 2, Appli
5	5810	98.1	1192	9	US-09-893-348-23	Sequence 23, Appl
6	4296.5	72.5	1163	9	US-09-893-348-18	Sequence 18, Appl
7	1495.5	25.2	373	9	US-09-789-386-6	Sequence 6, Appli
8	1495.5	25.2	373	9	US-09-765-205-6	Sequence 6, Appli
9	1495.5	25.2	373	9	US-09-893-348-24	Sequence 24, Appl
10	1495.5	275X7	373	15	US-10-060-036-72 LF	Sequence 72, Appl
11	1417	23.9	289	9	US-09-789-386-4	Sequence 4, Appli
12	1225.5	20.7	379	12	US-10-205-194-164	Sequence 164, App
13	1187	20.0	360	9	US-09-893-348-20	Sequence 20, Appl
14	931	15.7	199	9	US-09-893-348-25	Sequence 25, Appl
15	908	15.3	199	9	US-09-893-348-21	Sequence 21, Appl
16	791	13.4	777	12	US-10-205-219-93	Sequence 93, Appl
17	704	11.9	593	12	US-10-108-260A-2892	Sequence 2892, Ap
18	671	11.3	267	12	US-10-205-194-127	Sequence 127, App
19	625.5	10.6	236	9	US-09-729-674-20	Sequence 20, Appl
20	625.5	10.6	236	9	US-09-765-205-26075X7	Sequence 26075X7
21	625.5	10.6	269	15	US-10-106-698-6222	Sequence 6222, Ap
22	541.5	9.1	168	11	US-09-809-391-563	Sequence 563, App
23	541.5	9.1	168	12	US-09-882-171-563	Sequence 563, App
24	494	8.3	118	12	US-10-264-237-1568	Sequence 1568, Ap
25	340.5	5.7	1095	15	US-10-128-714-8305	Sequence 8305, Ap
26	331	5.6	66	9	US-09-758-140-20	Sequence 20, Appl
27	331	5.6	66	9	US-09-972-599A-20	Sequence 20, Appl
28	330	5.6	66	9	US-09-972-599A-22	Sequence 22, Appl
29	330	5.6	66	11	US-09-972-546-7	Sequence 7, Appli
30	311.5	5.3	98	15	US-10-050-704-179	Sequence 179, App
31	285	4.8	92	11	US-09-809-391-411	Sequence 411, App
32	285	4.8	92	12	US-09-882-171-411	Sequence 411, App
33	283.5	4.8	161	9	US-09-925-302-808	Sequence 808, App
34	276.5	4.7	1786	10	US-09-742-096-3	Sequence 3, Appli
35	275	4.6	1000	15	US-10-128-714-3305	Sequence 3305, Ap
36	274	4.6	3507	12	US-10-369-493-5784	Sequence 5784, Ap
37	273.5	4.6	2665	9	US-09-864-761-34248	Sequence 34248, A
38	273.5	4.6	3664	15	US-10-177-293-423	Sequence 423, App
39	267.5	4.5	1601	10	US-09-862-027-40	Sequence 40, Appl
40	265	4.5	1596	10	US-09-902-432-4	Sequence 4, Appli
41	263.5	4.4	2409	15	US-10-177-293-90	Sequence 90, Appl
42	262.5	4.4	6642	12	US-10-369-493-5013	Sequence 5013, Ap
43	258	4.4	1616	11	US-09-820-843A-16	Sequence 16, Appl
44	258	4.4	3913	12	US-10-334-143-45	Sequence 45, Appl
45	257.5	4.3	1781	9	US-09-738-877-3	Sequence 3, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-758-140-6

; Sequence 6, Application US/09758140

; Patent No. US20020012965A1

; GENERAL INFORMATION:

; APPLICANT: Strittmatter, Stephen M.



Db 481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540  
 Qy 527 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600  
 Qy 587 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 646  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 660  
 Qy 647 EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 705  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720  
 Qy 706 MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI 765  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI 780  
 Qy 766 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTAVY 825  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTAVY 840  
 Qy 826 SNDDLFISKEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 885  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 841 SNDDLFISKEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900  
 Qy 886 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVSALGH 945  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVSALA- 959  
 Qy 946 TQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSA DLGKTSVVDLLYWRDIKKTGV 1005  
 |||||||||||||||| : ||||||||||||||||||||  
 Db 960 TQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV 1019  
 Qy 1006 VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES 1065  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1020 VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES 1079  
 Qy 1066 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLT 1125  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1080 EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLT 1139  
 Qy 1126 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1140 LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 2

US-09-972-599A-6

; Sequence 6, Application US/09972599A

; Patent No. US20020077295A1

; GENERAL INFORMATION:

; APPLICANT: STRITTMATTER, STEPHEN M.

; TITLE OF INVENTION: NOGO-RECEPTOR-MEDIATED BLOCKADE OF AXONAL GROWTH

; FILE REFERENCE: C077 CIP US

; CURRENT APPLICATION NUMBER: US/09/972,599A

; CURRENT FILING DATE: 2001-10-06

; PRIOR APPLICATION NUMBER: PCT/US01/01041  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 09/758,140  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 60/236,378  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/207,366  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/175,707  
; PRIOR FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-972-599A-6

Query Match 98.2%; Score 5815; DB 9; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 3.5e-284;  
Matches 1160; Conservative 4; Mismatches 13; Indels 16; Gaps 3;

```
Qy      1 MEDLDQSPLVSSSDSPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:      :      : |||
Db    181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy    227 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy    287 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV 346
      |||
Db    301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDEVV 360

Qy    347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406
      |||
Db    361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy    407 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466
      |||
Db    421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy    467 PLLDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 526
      |||
Db    481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTE 540
```

Qy	527	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF	586
Db	541	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF	600
Qy	587	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDFSDYSE	705
Db	661	EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDFSDYSE	720
Qy	706	MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	765
Db	721	MAKVEQPVPDHSSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	840
Qy	826	SNDDLFIKSEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI	885
Db	841	SNDDLFIKSEAQIRETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI	900
Qy	886	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVSALGH	945
Db	901	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVSALA-	959
Qy	946	TQAEIESIVKPKVLEKEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGV	1005
Db	960	TQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSALSSTSVVDLLYWRDIKKTGV	1019
Qy	1006	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1065
Db	1020	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1079
Qy	1066	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT	1125
Db	1080	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLT	1139
Qy	1126	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	1140	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

RESULT 3

US-10-060-036-71

; Sequence 71, Application US/10060036  
 ; Publication No. US20030073144A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Benson, Darin R.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Persing, David H.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Jiang, Yuqiu



; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060,036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 71  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-060-036-71

Query Match 98.2%; Score 5815; DB 15; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 3.5e-284;  
Matches 1160; Conservative 4; Mismatches 13; Indels 16; Gaps 3;

```
Qy      1 MEDLDQSPLVSSSDSPPRQPAPFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPPRQPAPFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSSTVPAP 120

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGA-----VXKXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:
Db    181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy    227 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy    287 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHXQQELPTALTCLKVKEDEVV 346
      |||
Db    301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKEDEVV 360

Qy    347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406
      |||
Db    361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy    407 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466
      |||
Db    421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy    467 PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVTE 526
      |||
Db    481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy    527 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586
      |||
Db    541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
```

Qy	587	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	705
Db	661	EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	720
Qy	706	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI	765
Db	721	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY	840
Qy	826	SNDDLFISKEAQIRETETFSDDSPIEIIDFPTLISSKTD SFSKLAREYTDLEVSHKSEI	885
Db	841	SNDDLFISKEAQIRETETFSDDSPIEIIDFPTLISSKTD SFSKLAREYTDLEVSHKSEI	900
Qy	886	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS SKVLLLPPDV SALGH	945
Db	901	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS SKVLLLPPDV SALA-	959
Qy	946	TQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV	1005
Db	960	TQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV	1019
Qy	1006	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES	1065
Db	1020	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLES	1079
Qy	1066	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1125
Db	1080	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLT	1139
Qy	1126	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKI QAKI PGLKRKAE	1178
Db	1140	LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKI QAKI PGLKRKAE	1192

RESULT 4

US-09-789-386-2

; Sequence 2, Application US/09789386  
 ; Patent No. US20020010324A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MICHALOVICH, DAVID  
 ; APPLICANT: PRINJHA, RABINDER KUMAR  
 ; TITLE OF INVENTION: NOVEL COMPOUNDS  
 ; FILE REFERENCE: GP-30165-C1  
 ; CURRENT APPLICATION NUMBER: US/09/789,386  
 ; CURRENT FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: U.K. 9916898.1  
 ; PRIOR FILING DATE: 1999-07-19  
 ; PRIOR APPLICATION NUMBER: U.K. 9816024.5  
 ; PRIOR FILING DATE: 1998-07-22  
 ; PRIOR APPLICATION NUMBER: US 09/359,208



Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	705
Db	661	EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	720
Qy	706	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMI	765
Db	721	MAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMI	780
Qy	766	EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	825
Db	781	EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	840
Qy	826	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLI SSKTDSFSKLAREYTDLEVSHKSEI	885
Db	841	SNDDLFI SKEAQIRETETFS DSSPIEI IDEFPTLI SSKTDSFSKLAREYTDLEVSHKSEI	900
Qy	886	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI SFSDDFS KNGSATS KVVLLLPDVSALGH	945
Db	901	ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI SFSDDFS KNGSATS KVVLLLPDVSALA-	959
Qy	946	TQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDI KKTGV	1005
Db	960	TQAEIESIVKPKVLVKEAEKKLP SDTEKEDRSPSAIF SAELSKTSVVDLLYWRDI KKTGV	1019
Qy	1006	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1065
Db	1020	VFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLES	1079
Qy	1066	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLT	1125
Db	1080	EVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGALFNGLT	1139
Qy	1126	LLILALISLFSVPVIERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	1140	LLILALISLFSVPVIERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

# RESULT 5

US-09-893-348-23

; Sequence 23, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

; FILE REFERENCE: EIS-SCHWARTZ=2A

; CURRENT APPLICATION NUMBER: US/09/893,348

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/314,161

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: US 09/218,277

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: PCT/US98/14715

; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: IL 124500  
; PRIOR FILING DATE: 1998-05-19  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-893-348-23

Query Match 98.1%; Score 5810; DB 9; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 6.2e-284;  
Matches 1159; Conservative 4; Mismatches 14; Indels 16; Gaps 3;

```
Qy      1 MEDLDQSPLVSSSDSPPRQPAPFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPPRQPAPFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP 226
      |||:
Db    181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy    227 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 286
      |||
Db    241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy    287 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV 346
      |||
Db    301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNNQQELPTALTCLKVKEDEVV 360

Qy    347 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 406
      |||
Db    361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy    407 DKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 466
      |||
Db    421 DKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy    467 PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTKVTE 526
      |||
Db    481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLLVAAQDSETDYVTTDNLTKVTE 540

Qy    527 EVVANMPEGLTPDLVQFACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 586
      |||
Db    541 EVVANMPEGLTPDLVQFACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy    587 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE 646
      |||
```



```

; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-893-348-18

```

```

Query Match          72.5%; Score 4296.5; DB 9; Length 1163;
Best Local Similarity 74.0%; Pred. No. 7.4e-208;
Matches 885; Conservative 104; Mismatches 156; Indels 51; Gaps 19;

```

```

Qy      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELVLERK 58
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||||
Db      1 MEDIDQSSSVSSSTDSPRPAPPAFKYQFVTEPEDEEEDDEEEDDEDLEELVLERK 60

Qy     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSTVP 118
      ||||| || ||| |||:| | :| ||||| ||||| |||||:| | : |
Db     61 PAAGLSAAAVP--PAAAAPLLDSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115

Qy    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178
      ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166

Qy    179 RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSL 224
      ||| : |||| |||||:|:||||||| |||||
Db    167 RGSGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL 226

Qy    225 SPLSAAAFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS 284
      ||| |||| ||||| | :|||:| :|:|||||: |:| :| |||||
Db    227 SPLSTVSFKEHGYLGNLSAVSSSEGTTIETLNEASKELPERATNPFVNRDLAEFSELEYS 286

Qy    285 EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDE 344
      ||||| ||| |||:| | :|:|:|:|:|:| | : || ||| |||
Db    287 EMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KEDR 340

Qy    345 VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE 403
      ||| || | ||| :|:| ||:|||||||: ||||: | | |:| | :|:|
Db    341 VVSPEKTMDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANVE 396

Qy    404 SKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIAT 463
      |||:| | |||| : |||| |:| ||||| :| | ||||| | :|||
Db    397 SKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA 455

Qy    464 NIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTK 523
      | ||||| |||| |||||:|:|:| | ||||| || |||| ||||| |:|
Db    456 NTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLTK 514

Qy    524 VTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLC 583
      ||| |:| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    515 VTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLC 574

```





```
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-789-386-6
```

```
Query Match          25.2%; Score 1495.5; DB 9; Length 373;
Best Local Similarity 31.6%; Pred. No. 1.2e-67;
Matches 372; Conservative 1; Mismatches 0; Indels 805; Gaps 1;
```

```
Qy      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP 120

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAAFKEHEYLG 240
      |||:|
Db    181 SSGSV----- 185

Qy    241 LSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVSPKAESAV 300
Db    186 ----- 185

Qy    301 IVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDDEVVSSEKAKDSFNEKRV 360
Db    186 ----- 185

Qy    361 AVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNH 420
Db    186 ----- 185

Qy    421 EKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDE 480
Db    186 ----- 185

Qy    481 KKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDL 540
Db    186 ----- 185

Qy    541 VQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDI 600
Db    186 ----- 185

Qy    601 VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIKE 660
Db    186 ----- 185
```

Qy 661 EIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELV 720

Db 186 ----- 185

Qy 721 EDSSPDSEPVDLFDSDSI PDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEG 780

Db 186 ----- 185

Qy 781 GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQIRE 840

Db 186 ----- 185

Qy 841 TETFSDDSSPIEII DEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL 900

Db 186 ----- 185

Qy 901 PHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVVLLPPDV SALGHTQAEIESIVKPKVLE 960

Db 186 ----- 185

Qy 961 KEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF 1020

Db 186 ----- VVDLLYWRDIKKTGVVFGASLFLLLSLTVF 215

Qy 1021 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 1080

Db 216 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 275

Qy 1081 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 1140

Db 276 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 335

Qy 1141 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178

Db 336 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 373

# RESULT 8

US-09-765-205-6

; Sequence 6, Application US/09765205

; Patent No. US20020034800A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Li

; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES

; FILE REFERENCE: 1458.004/200130.449

; CURRENT APPLICATION NUMBER: US/09/765,205

; CURRENT FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: US/09/212,440

; PRIOR FILING DATE: 1998-12-16

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 6

; LENGTH: 373

; TYPE: PRT

; ORGANISM: human

US-09-765-205-6



Db 186 ----- 185  
 Qy 841 TETFS DSSPIEI IDEFPTLISSKTD SFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL 900  
 Db 186 ----- 185  
 Qy 901 PHDLSLKNIQPKVEEKISFSDDFS KN GSATSKVLLLPDVSALGHTQAEIESIVKPKVLE 960  
 Db 186 ----- 185  
 Qy 961 KEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF 1020  
 Db 186 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVF 215  
 V Qy 1 1 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 1080  
 Db 216 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 275  
 Qy 1081 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 1140  
 Db 276 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 335  
 Qy 1141 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 Db 336 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 373

RESULT 9

US-09-893-348-24

; Sequence 24, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, G

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

; FILE REFERENCE: EIS-SCHWARTZ=2A

; CURRENT APPLICATION NUMBER: US/09/893,348

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/314,161

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: US 09/218,277

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: PCT/US98/14715

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: IL 124500

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 24

; LENGTH: 373

; TYPE: PRT

; ORGANISM: Homo sapiens

50XL

US-09-893-348-24

Query Match 25.2%; Score 1495.5; DB 9; Length 373;  
Best Local Similarity 31.6%; Pred. No. 1.2e-67;  
Matches 372; Conservative 1; Mismatches 0; Indels 805; Gaps 1;

```
Qy      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
      |||
Db      1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
      |||
Db     61 AGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKRRG 180
      |||
Db    121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKRRG 180

Qy    181 SSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAAASFKEHEYLGN 240
      |||
Db    181 SSGSV----- 185

Qy    241 LSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFSVSPKAESAV 300
Db    186 ----- 185

Qy    301 IVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEVVSSEKAKDSFNEKRV 360
Db    186 ----- 185

Qy    361 AVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNH 420
Db    186 ----- 185

Qy    421 EKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTDE 480
Db    186 ----- 185

Qy    481 KKIEEKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDL 540
Db    186 ----- 185

Qy    541 VQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDI 600
Db    186 ----- 185

Qy    601 VMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGIKE 660
Db    186 ----- 185

Qy    661 EIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDHSELV 720
Db    186 ----- 185

Qy    721 EDSSPDSEPVDLFSDDISPDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEG 780
Db    186 ----- 185
```

```

Qy      781 GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTAVYSNDDLFIKSEAQIRE 840
Db      186 ----- 185
Qy      841 TETFSOSSPIEIIDEFPTLISSKTDSEFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL 900
Db      186 ----- 185
Qy      901 PHDLSLKNIQPKVEEKISFSDDFSKNKSATSKVLLLPPDVSA LGHTQAEIESIVKPKVLE 960
Db      186 ----- 185
Qy      961 KEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF 1020
Db      186 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVF 215
Qy     1021 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNS 1080
Db     216 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAYLESEVAISEELVQKYSNS 275
Qy     1081 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 1140
Db     276 ALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 335
Qy     1141 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178
Db     336 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 373

```

# RESULT 10

US-10-060-036-72

; Sequence 72, Application US/10060036

; Publication No. US20030073144A1

; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.

; AP : Lodes, Michael J.

LERKP45

; APPLICANT: Persing, David H.

; APPLICANT: Hepler, William T.

; APPLICANT: Jiang, Yuqiu

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER

; FILE REFERENCE: 210121.566

; CURRENT APPLICATION NUMBER: US/10/060,036

; CURRENT FILING DATE: 2002-01-30

; NUMBER OF SEQ ID NOS: 4560

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 72

; LENGTH: 373

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-060-036-72

Query Match 25.2%; Score 1495.5; DB 15; Length 373;

Best Local Similarity 31.6%; Pred. No. 1.2e-67;

Matches 372; Conservative 1; Mismatches 0; Indels 805; Gaps 1;

Qy	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDLEELVLERKPA	60
Db	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDLEELVLERKPA	60
Qy	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSTVPAP	120
Qy	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Db	121	SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPAEPVWTPPAPAPAAPPSTPAAPKRRG	180
Qy	181	SSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAASFKEHEYLGN	240
Db	181	SSGSV-----	18950XA
Qy	241	LSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFSVSPKAESAV	300
Db	186	-----	185
Qy	301	IVANPREEIIVKNKDDEEEKLVSNNILHXQQELPTALTCLVKEDDEVSSSEKAKDSFNEKRV	360
Db	186	-----	185
Qy	361	AVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTNH	420
Db	186	-----	185
Qy	421	EKDESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFLLEDPTSSENXTDE	480
Db	186	-----	185
Qy	481	KKIEEKKAQIVTEKNTSTKTNSNPFFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTPDL	540
Db	186	-----	185
Qy	541	VQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFESEATPSPVLPDI	600
Db	186	-----	185
Qy	601	VMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEMSLSLKVSIGIKE	660
Db	186	-----	185
Qy	661	EIKEPENINAALQETEAPYISIACDLIKETKLSEAEPAPDFS DYSEMAKVEQVPDHSELV	720
Db	186	-----	185
Qy	721	EDSSPDSEPVDLFSDDSIPDV PQQDET VMLVKESLTETS FESMI EYENKEKLSALPPEG	780
Db	186	-----	185
Qy	781	GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIP LQMEELS TAVYSNDDL FISKEAQIRE	840
Db	186	-----	185

Qy 841 TETFSDDSSPIEIIDFPTLISSKTD SFSK LAREYTDLEVSHKSEIANAPDGAGSLPCTEL 900  
 Db 186 ----- 185  
 Qy 901 PHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SALGHTQAEIESIVKPKVLE 960  
 Db 186 ----- 185  
 Qy 961 KEAEKKLPSTDEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF 1020  
 Db 186 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVF 215  
 Qy 1021 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 1080  
 Db 216 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 275  
 Qy 1081 ALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 1140  
 Db 276 ALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 335  
 Qy 1141 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178  
 Db 336 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 373

RESULT 11

US-09-789-386-4

; Sequence 4, Application US/09789386  
 ; Patent No. US20020010324A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MICHALOVICH, DAVID  
 ; APPLICANT: PRINJHA, RABINDER KUMAR  
 ; TITLE OF INVENTION: NOVEL COMPOUNDS  
 ; FILE REFERENCE: GP-30165-C1  
 ; CURRENT APPLICATION NUMBER: US/09/789,386  
 ; CURRENT FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: U.K. 9916898.1  
 ; PRIOR FILING DATE: 1999-07-19  
 ; PRIOR APPLICATION NUMBER: U.K. 9816024.5  
 ; PRIOR FILING DATE: 1998-07-22  
 ; PRIOR APPLICATION NUMBER: US 09/359,208  
 ; PRIOR FILING DATE: 1999-07-22  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: HOMO SAPIENS  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (31)(138)

US-09-789-386-4

Query Match 23.9%; Score 1417; DB 9; Length 289;  
 Best Local Similarity 99.3%; Pred. No. 7.6e-64;  
 Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



Qy 193 MDLKEQPGNTISAGQEDFPSVLLETAASXPSSLPLSAASFKEHEYLGNLSTVLPTEGTLQ 252  
 |||  
 Db 3 MDLKEQPGNTISAGQEDFPSVLLETAASXPSSLPLSAASFKEHEYLGNLSTVLPTEGTLQ 62

Qy 253 ENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFVSVPKAE SAVIVANPREEIIVK 312  
 |||  
 Db 63 ENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSFVSVPKAE SAVIVANPREEIIVK 122

Qy 313 NKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYAD 372  
 |||  
 Db 123 NKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYAD 182

Qy 373 FKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDESSNDDTS 432  
 |||  
 Db 183 FKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCFADSLEQTNHEKDESSNDDTS 242

Qy 433 FPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXTD 479  
 |||  
 Db 243 FPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENKTD 289

RESULT 12

US-10-205-194-164

; Sequence 164, Application US/10205194  
 ; Publication No. US20030134301A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Warner-Lambert Company  
 ; APPLICANT: Lee, Kevin  
 ; APPLICANT: Dixon, Alistair  
 ; APPLICANT: Brooksbank, Robert  
 ; APPLICANT: Pinnock, Robert  
 ; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
 ; FILE REFERENCE: WL-A-018201  
 ; CURRENT APPLICATION NUMBER: US/10/205,194  
 ; CURRENT FILING DATE: 5200-07-24  
 ; PRIOR APPLICATION NUMBER: GB 0118354.0  
 ; PRIOR FILING DATE: 2001-07-27  
 ; NUMBER OF SEQ ID NOS: 177  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 164  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Foocen-m2 reticulon  
 US-10-205-194-164

Query Match 20.7%; Score 1225.5; DB 12; Length 379;  
 Best Local Similarity 28.6%; Pred. No. 4.6e-54;  
 Matches 338; Conservative 17; Mismatches 22; Indels 803; Gaps 10;

Qy 1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58  
 |||: ||| |||| |||| |||| |||| |||| ||||: ||: |||| |||| ||||  
 Db 1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEDEDLEELEVLERK 60

Qy 59 PAAGLSAAPVPTAPAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVP 118  
 ||||| || ||| |||: ||: |||| |||| |||| |||| ||||: ||: ||

Db	61	PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Qy	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKR	178
Db	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Qy	179	RGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAASFKEHEYL	238
Db	167	RGS-----	169
Qy	239	GNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMSSFSVSPKAES	298
Db	170	-----	169
Qy	299	AVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTKLVKEDEVVSSEKAKDSFNEK	358
Db	170	-----	169
Qy	359	RVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQT	418
Db	170	-----	169
Qy	419	NHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXT	478
		:	
Db	170	-----GSVDETLF-----	177
Qy	479	DEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTP	538
Db	178	-----	177
Qy	539	DLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFESESEATPSPVLP	598
Db	178	-----	177
Qy	599	DIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGI	658
		:   :	
Db	178	-----ALPAA-----	182
Qy	659	KEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSE	718
Db	183	-----	182
Qy	719	LVEDSSPDSEPVDLFSDDSI PDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPP	778
Db	183	-----SEPV-----	186
Qy	779	EGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFI SKEAQI	838
Db	187	-----	186
Qy	839	RETETFSDDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCT	898
Db	187	-----	186
Qy	899	ELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLLPPDV SALGHTQAEIESIVKPKV	958
Db	187	-----	186

```

Qy      959 LEKEAEKKLPSDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLT 1018
          :||
Db      187 -----IPS-----SAVVDLLYWRDIKKTGVVFGASLFLLLSLT 219

Qy      1019 VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS 1078
          |||
Db      220 VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS 279

Qy      1079 NSALGHVNCTIKELRRLFLVDDLVDLKFVFLMWVFTYVGALFNGLTLLILALISLFSVP 1138
          |||
Db      280 NSALGHVNSTIKELRRLFLVDDLVDLKFVFLMWVFTYVGALFNGLTLLILALISLFSIP 339

Qy      1139 VIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178
          |||
Db      340 VIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 379

```

# RESULT 13

US-09-893-348-20

; Sequence 20, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

; FILE REFERENCE: EIS-SCHWARTZ=2A

; CURRENT APPLICATION NUMBER: US/09/893,348

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/314,161

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: US 09/218,277

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: PCT/US98/14715

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: IL 124500

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20

; LENGTH: 360

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-893-348-20

Query Match 20.0%; Score 1187; DB 9; Length 360;

Best Local Similarity 27.7%; Pred. No. 3.7e-52;

Matches 327; Conservative 12; Mismatches 19; Indels 822; Gaps 7;

```

Qy      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58
          |||
Db      1 MEDIDQSSLVSSSTDSPPPPPAFKYQFVTEPEDEEDEEEEEEEEEDEDEDLEELEVLERK 60

```

QY	59	PAAGLSAAPVPTAPAAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVP	118
Db	61	PAAGLSAAAVP--PAAAAPLLDFFSSDSVPPAPRGPLPAAPPAAAPERQPSWERSPAA--P	115
QY	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Db	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
QY	179	RGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSSLSPLSAASFKEHEYL	238
Db	167	RGS-----	169
QY	239	GNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSVSPKAES	298
Db	1700	-----	169
QY	299	AVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDEVVSSEKAKDSFNEK	358
Db	170	-----	169
QY	359	RVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQT	418
Db	170	-----	169
QY	419	NHEKDESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLEDPTSENXT	478
Db	170	-----	169
QY	479	DEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEEVVANMPEGLTP	538
Db	170	-----	169
QY	539	DLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFESESEATPSPVLP	598
Db	170	-----	169
QY	599	DIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYEEAMSVSLKVSGI	658
Db	170	-----	169
QY	659	KEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPDPHSE	718
Db	170	-----	169
QY	719	LVEDSSPDSEPVDLFSDDSIIPDVPQKQDETVMVLKESLTETSFESMIEYENKEKLSALPP	778
Db	170	-----	169
QY	779	EGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLOMEELSTAVYSNDDLFIKEAQI	838
Db	170	-----	169
QY	839	RETETFSDDSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCT	898
Db	170	-----	172
QY	899	ELPHDLSLKNIQPKVEEKISFSDDFSKNKSATSKVLLLPDPVVSALGHTOAEIESIVKPKV	958

```

Db      173 ----- 172
Qy      959 LEKEAEKKLPSTEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLT 1018
          |||
Db      173 -----VVDLLYWRDIKKTGVVFGASLFLLLSLT 200
Qy      1019 VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS 1078
          |||
Db      201 VFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYS 260
Qy      1079 NSALGHVNCTIKELRRLFLVDDLVDLKFVLMWVFTYVGALFNGLTLLILALISLFSVP 1138
          |||
Db      261 NSALGHVNSTIKELRRLFLVDDLVDLKFVLMWVFTYVGALFNGLTLLILALISLFSIP 320
Qy      WO2001393Y 1.5X12DHXLGLANKNVKDAMAKIQAQIPGLKRKAE 1178
          |||
Db      321 VIYERHQVQIDHYLGLANKSVKDAMAKIQAQIPGLKRKAD 360

```

RESULT 14

US-09-893-348-25

; Sequence 25, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

; FILE REFERENCE: EIS-SCHWARTZ=2A

; CURRENT APPLICATION NUMBER: US/09/893,348

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/314,161

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: US 09/218,277

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: PCT/US98/14715

; PRIOR FILING DATE: 1998-07-21

; Pn LOCATION NUMBER: IL 124500

otide30 d

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 199

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-893-348-25

Query Match 15.7%; Score 931; DB 9; Length 199;

Best Local Similarity 99.0%; Pred. No. 1.3e-39;

Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 1047
          | |||

```

```

Db          9 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGUI 68
Qy         1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKF 1107
           |||
Db          69 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKF 128
Qy         1108 AVLMMWFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167
           |||
Db          129 AVLMMWFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 188
Qy         1168 AKIPGLKRKAE 1178
           |||
Db          189 AKIPGLKRKAE 199

```

RESULT 15

US-09-893-348-21

; Sequence 21, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila

; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES

; FILE REFERENCE: EIS-SCHWARTZ=2A

; CURRENT APPLICATION NUMBER: US/09/893,348

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/314,161

; PRIOR FILING DATE: 1999-05-19

; PRIOR APPLICATION NUMBER: US 09/218,277

; PRIOR FILING DATE: 1998-12-22

; PRIOR APPLICATION NUMBER: PCT/US98/14715

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: IL 124500

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 21

; LENGTH: 199

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-893-348-21

Query Match 15.3%; Score 908; DB 9; Length 199;

Best Local Similarity 96.3%; Pred. No. 1.9e-38;

Matches 184; Conserved 3; Mismatch 4; Indels 0; Gaps 0;

```

Qy          988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGUI 1047
           |||
Db           9 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGUI 68
Qy         1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKF 1107
           |||
Db          69 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKF 128

```

```

Qy      1108 AVLMWVFTYVVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167
          ||||||||||||||||||||||||||||:||||||| |||||||||||:|||||||
Db      129  AVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ 188

Qy      1168 AKIPGLKRKAE 1178
          |||||||||:
Db      189  AKIPGLKRKAD 199

```

Search completed: January 22, 2004, 16:44:57  
Job time : 63.0516 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 16:31:15 ; Search time 19.6482 Seconds  
 (without alignments)  
 2819.465 Million cell updates/sec

Title: US-09-830-972-29  
 Perfect score: 5923  
 Sequence: 1 MEDLDQSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRKAE 1178

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5815	98.2	1192	1	RTN4_HUMAN	Q9nqc3 homo sapien
2	4296.5	72.5	1163	1	RTN4_RAT	Q9jk11 rattus norv
3	915	15.4	199	1	RTN4_MOUSE	Q99p72 mus musculu
4	791	13.4	777	1	RTN1_RAT	Q64548 rattus norv
5	787.5	13.3	776	1	RTN1_HUMAN	Q16799 homo sapien
6	626	10.6	237	1	RTNp1E008E52	Q9es97 mus musculu
7	625.5	10.6	236	1	RTN3_HUMAN	O95197 homo sapien
8	514	8.7	545	1	RTN2_HUMAN	O75298 homo sapien
9	500	8.4	471	1	RTN2_MOUSE	O70622 mus musculu
10	324.5	5.5	865	1	CPN_DROME	Q02910 drosophila
11	320	5.4	5147	1	PCLO_HUMAN	Q9y6v0 homo sapien
12	315	5.3	5038	1	PCLO_MOUSE	Q9qyx7 mus musculu
13	299	5.0	3924	1	ANK2_HUMAN	Q01484 homo sapien
14	296.5	5.0	2468	1	MAPB_HUMAN	P46821 homo sapien
15	289	4.9	2459	1	MAPB_RAT	P15205 rattus norv
16	289	4.9	2464	1	MAPB_MOUSE	P14873 mus musculu
17	287	4.8	5120	1	PCLO_CHICK	Q9pu36 gallus gall



18	276	4.7	5085	1	PCLO_RAT	Q9jks6	rattus norv
19	264.5	4.5	8545	1	ANC1_CAEEL	Q9n4m4	caenorhabdi
20	263.5	4.4	3396	1	PGCV_HUMAN	P13611	homo sapien
21	262.5	4.4	6632	1	UN89_CAEEL	O01761	caenorhabdi
22	260.5	4.4	3381	1	PGCV_BOVIN	P81282	bos taurus
23	259.5	4.4	1828	1	MAP2_MOUSE	P20357	mus musculu
24	258.5	4.4	2805	1	MAPA_HUMAN	P78559	homo sapien
25	258	4.4	1616	1	P200_MYCGE	Q49429	mycoplasma
26	257.5	4.3	1781	1	AK12_HUMAN	Q02952	homo sapien
27	257.5	4.3	2738	1	PGCV_RAT	Q9erb4	rattus norv
28	253	4.3	1189	1	YJH6_YEAST	P47035	saccharomyc
29	252	4.3	4377	1	ANK3_HUMAN	Q12955	homo sapien
30	250.5	4.2	1861	1	MAP2_RAT	P15146	rattus norv
31	249.5	4.2	1827	1	MAP2_HUMAN	P11137	homo sapien
32	248	4.2	3421	1	TEGU_HSVEB	P28955	equine herp
33	244	4.1	1362	1	BRD4_HUMAN	O60885	homo sapien
34	242	4.1	1790	1	USO1_YEAST	P25386	saccharomyc
35	239	4.0	1744	1	TANA_XENLA	Q01550	xenopus lae
36	238.5	4.0	1411	1	TCOF_HUMAN	Q13428	homo sapien
37	236.5	4.0	3358	1	PGCV_MOUSE	Q62059	mus musculu
38	235.5	4.0	1723	1	AIM1_HUMAN	Q9y4k1	homo sapien
39	233.5	3.9	1087	1	NFH_MOUSE	P19246	mus musculu
40	232	3.9	1140	1	YM96_YEAST	Q04893	saccharomyc
41	131	3.9	1140	1	KI67_HUMAN	P46013	homo sapien
42	229.5	3.9	1367	1	AMYH_YEAST	P08640	saccharomyc
43	228.5	3.9	1306	1	MSB2_YEAST	P32334	saccharomyc
44	226.5	3.8	1157	1	BBC1_YEAST	P47068	saccharomyc
45	226	3.8	1379	1	YFF9_SCHPO	O14066	schizosacch

# ALIGNMENTS

## RESULT 1

### RTN4\_HUMAN

ID RTN4\_HUMAN STANDARD; PRT; 1192 AA.  
AC Q9NQC3; Q9A962; Q9BXG5; Q9H212; Q9H3I3; Q9UQ42; Q9Y293; Q9Y2Y7;  
AC Q9Y5U6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)  
DE (Neuroendocrine-specific protein) (NSP) (Neuroendocrine-specific  
DE protein C homolog) (RTN-  
GN RTN4 OR NOGO OR ASY OR KIAA0886.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=20129242; PubMed=10667780;  
RA Prinjha R., Moore S.E., Vinson M., Blake S., Morrow R., Christie G.,  
RA Michalovich D., Simmons D.L., Walsh F.S.;  
RT "Inhibitor of neurite outgrowth in humans."  
RL Nature 403:383-384(2000).  
RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=21010696; PubMed=11126360;  
 RA Tagami S., Eguchi Y., Kinoshita M., Takeda M., Tsujimoto Y.;  
 RT "A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on  
 RT endoplasmic reticulum and reduces their anti-apoptotic activity.";   
 RL Oncogene 19:5736-5746(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RX MEDLINE=20237542; PubMed=10773680;  
 RA Yang J., Yu L., Bi A.D., Zhao S.-Y.;  
 RT "Assignment of the human reticulon 4 gene (RTN4) to chromosome  
 RT 2p14-->2p13 by radiation hybrid mapping.";   
 RL Cytogenet. Cell Genet. 88:101-102(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 4).  
 RA Jin W.-L., Ju G.;  
 RT "Developmentally-regulated alternative splicing in a novel Nogo-A.";   
 RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).  
 RC TISSUE=Placenta, and Skeletal muscle;  
 RA Ito T., Schwartz S.M.;  
 RT "Cloning of a member of the reticulon gene family in human.";   
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Fibroblast;  
 RA Yutsudo M.;  
 RT "Isolation of a cell death-inducing gene.";   
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Pituitary;  
 RA Song H., Peng Y., Zhou J., Huang Q., Dai M., Mao Y.M., Yu Y., Xu X.,  
 RA Luo B., Hu R., Chen J.;  
 RT "Human neuroendocrine-specific protein C (NSP) homolog gene.";   
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,  
 RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,  
 RA Yu J., Han L.H.;  
 RT "Novel human cDNA clone with function of inhibiting cancer cell  
 RT growth.";   
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=99156230; PubMed=10048485;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";   
 RL DNA Res. 5:355-364(1998).  
 RN [10]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).  
 RC TISSUE=Brain, Pancreas, Placenta, and Skeletal muscle;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., rgren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [11]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,  
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,  
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;  
 RT "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560(2000).  
 RN [12]  
 RP SEQUENCE OF 482-1192 FROM N.A. (ISOFORM 1/4).  
 RC TISSUE=Brain;  
 RA Mao Y.M., Xie Y., Zheng Z.H.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [13]  
 RP SEQUENCE OF 186-1192 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Testis;  
 RA Sha J.H., Zhou Z.M., Li J.M.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [14]  
 RP TOPOLOGY.  
 RC TISSUE=Brain;  
 RX MEDLINE=20129259; PubMed=10667797;  
 RA GrandPre T., Nakamura F., Vartanian T., Strittmatter S.M.;  
 RT "Identification of the Nogo inhibitor of axon regeneration as a  
 RT Reticulon protein.";  
 RL Nature 403:439-444(2000).  
 RN [15]  
 RP FUNCTION.  
 RC TISSUE=Brain;  
 RX =21069055; PubMed=11201742; --- 1412  
 RA Fournier A.E., Grandpre T., Strittmatter S.M.;  
 RT "Identification of a receptor mediating Nogo-66 inhibition of axonal

RT regeneration.";  
 RL Nature 409:341-346(2001).  
 RN [16]  
 RP REVIEW.  
 RX MEDLINE=21888956; PubMed=11891768;  
 RA Ng C.E.L., Tang B.L.;  
 RT "Nogos and the Nogo-66 receptor: factors inhibiting CNS neuron  
 RT regeneration.";  
 RL J. Neurosci. Res. 67:559-565(2002).  
 CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help  
 CC block the regeneration of the nervous central system in adults.  
 CC Isoform 2 reduces the anti-apoptotic activity of Bcl-xl and Bcl-2.  
 CC This is likely consecutive to their change in subcellular  
 CC location, from the mitochondria to the endoplasmic reticulum,  
 CC after binding and sequestration.  
 CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum. Anchored to the membrane of the endoplasmic reticulum  
 CC through 2 putative transmembrane domains.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1; Synonyms=RTN 4A, Nogo-A, RTN-xL;  
 CC IsoId=Q9NQC3-1; Sequence=Displayed;  
 CC Name=2; Synonyms=RTN 4B, Nogo-B, RTN-xS, Foocen-M;  
 CC IsoId=Q9NQC3-2; Sequence=VSP\_005655;  
 CC Name=3; Synonyms=RTN 4C, Nogo-C, Foocen-S;  
 CC IsoId=Q9NQC3-3; Sequence=VSP\_005652, VSP\_005653;  
 CC Name=4;  
 CC IsoId=Q9NQC3-4; Sequence=VSP\_005654;  
 CC -!- TISSUE SPECIFICITY: Isoform 1 is specifically expressed in brain  
 CC and testis and weakly in heart and skeletal muscle. Isoform 2 is  
 CC widely expressed excepted for the liver. Isoform 3 is expressed in  
 CC brain, skeletal muscle and adipocytes. Isoform 4 is testis-  
 CC specific.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -!- CAUTION: Ref.11 sequence differs from that shown due to  
 CC frameshifts in positions 1149 and 1156.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AJ251383; CAB99248.1; -.  
 DR EMBL; AJ251384; CAB99249.1; -.  
 DR EMBL; AJ251385; CAB99250.1; -.  
 DR EMBL; AB040462; BAB18927.1; -.  
 DR EMBL; AB040463; BAB18928.1; -.  
 DR EMBL; AF148537; AAG12176.1; -.  
 DR EMBL; AF148538; AAG12177.1; -.  
 DR EMBL; AF087901; AAG12205.1; -.  
 DR EMBL; AF320999; AAG40878.1; -.  
 DR EMBL; AF132047; AAD31021.1; -.  
 DR EMBL; AF132048; AAD31022.1; -.

Query Match 98.2%; Score 5815; DB 1; Length 1192;  
Best Local Similarity 97.2%; Pred. No. 2.2e-216;  
Matches 1160; Conservative 4; Mismatches 13; Indels 16; Gaps 3;

Qy	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Db	1	MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA	60
Qy	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
Db	61	AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVPAP	120
Qy	121	SPLSAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRRG	180
Db	121	SPLSAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPSTPAAPKRRG	180
Qy	181	SSGA-----VVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSP	226
Db	181	SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP	240
Qy	227	LSAASFKEHEYLGNLSTVLPTTEGLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	286
Db	241	LSAASFKEHEYLGNLSTVLPTTEGLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM	300
Qy	287	GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVKEDEVV	346
Db	301	GSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKEDEVV	360
Qy	347	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	406
Db	361	SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV	420
Qy	407	DKKCFADSLEQTNHEKDSSESSNDDTSFSPSTPEGIKDRSGAYITCAPFNPAATESIATNIF	466
Db	421	DKKCFADSLEQTNHEKDSSESSNDDTSFSPSTPEGIKDRSGAYITCAPFNPAATESIATNIF	480
Qy	467	PLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSETDYVTTDNLTKVTE	526
Db	481	PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE	540
Qy	527	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF	586
Db	541	EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF	600
Qy	587	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	646
Db	601	EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHEPENPPPYE	660
Qy	647	EAMSVSL-KVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE	705

Db	661		EAMSVSLKKVSGIKKEIKEPENINAALQTEAPYISACDLIKETKLSAEPAPDFSDYSE	720
Qy	706		MAKVEQPVDPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMI	765
Db	721		MAKVEQPVDPDHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMI	780
Qy	766		EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	825
Db	781		EYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVY	840
Qy	826		SNDDLFI SKEAQIRETETFS DSSPIEI IDEFP TLISSKTDSFSKLAREYTDLEVSHKSEI	885
Db	841		SNDDLFI SKEAQIRETETFS DSSPIEI IDEFP TLISSKTDSFSKLAREYTDLEVSHKSEI	900
Qy	886		ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SALGH	945
Db	901		ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATS KVL LPPDV SALA-	959
Qy	946		TQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIF SADLGKTSVVDLLYWRDIKKTGV	1005
Db	960		TQAEIESIVKPKVLVKEAEKKLP SDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGV	1019
Qy	1006		VFGASLFL LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLES	1065
Db	1020		VFGASLFL LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLES	1079
Qy	1066		EVAISEELVQKYSNSALGHVNCTIKELRRFLVDDLVDLSLKFVLMWVFTYVGALFNGLT	1125
Db	1080		EVAISEELVQKYSNSALGHVNCTIKELRRFLVDDLVDLSLKFVLMWVFTYVGALFNGLT	1139
Qy	1126		LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1178
Db	1140		LLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

RESULT 2

RTN4\_RAT

ID RTN4\_RAT STANDARD; PRT; 1163 AA.  
AC Q9JK11; Q9JK10; Q9R0D9; Q9WUE9; Q9WUF0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)  
DE (Glut4 vesicle 20 kDa protein).  
GN RTN4 OR NOGO.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.  
RC STRAIN=Sprague-Dawley; TISSUE=Adipocyte;  
RX MEDLINE=99249816; PubMed=10231557;  
RA Morris N.J., Ross S.A., Neveu J.M., Lane W.S., Lienhard G.E.;  
RT "Cloning and characterization of a 22 kDa protein from rat adipocytes:  
RT a new member of the reticulon family.";

RL Biochim. Biophys. Acta 1450:68-76(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RX MEDLINE=20129258; PubMed=10667796;  
 RA Chen M.S., Huber A.B., Van der Haar M.E., Frank M., Schnell L.,  
 RA Spillmann A.A., Christ F., Schwab M.E.;  
 RT "Nogo-A is a myelin-associated neurite outgrowth inhibitor and an  
 RT antigen for monoclonal antibody IN-1.";  
 RL Nature 403:434-439(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).  
 RC STRAIN=Wistar Kyoto; TISSUE=Vascular smooth muscle;  
 RA Ito T., Schwartz S.M.;  
 RT "Cloning of a member of the reticulon gene family in rat: one of two  
 RT minor splice variants.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE=22033691; PubMed=12037567;  
 RA GrandPre T., Li S., Strittmatter S.M.;  
 RT "Nogo-66 receptor antagonist peptide promotes axonal regeneration.";  
 RL Nature 417:547-551(2002).  
 CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help  
 CC block the regeneration of the nervous central system in adults (By  
 CC similarity).  
 CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-x1 and Bcl-2 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the  
 CC membrane of the endoplasmic reticulum through 2 putative  
 CC transmembrane domains (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1; Synonyms=Nogo-A, NI-220-250;  
 CC IsoId=Q9JK11-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Nogo-B, Foocen-M1;  
 CC IsoId=Q9JK11-2; Sequence=VSP\_005658;  
 CC Name=3; Synonyms=Nogo-C, VP20;  
 CC IsoId=Q9JK11-3; Sequence=VSP\_005656, VSP\_005657;  
 CC Name=4; Synonyms=Foocen-M2;  
 CC IsoId=Q9JK11-4; Sequence=VSP\_005659;  
 CC -!- TISSUE SPECIFICITY: Isoforms 1, 2 and 3 are present in optic  
 CC nerve, spinal cord and cerebral cortex. Isoforms 1 and 2 are  
 CC present in dorsal root ganglion, sciatic nerve and PC12 cells  
 CC after longer exposure. Isoforms 2 and 3 are detected in kidney,  
 CC cartilage, skin, lung and spleen. Isoform 3 is expressed at high  
 CC level in skeletal muscle. In adult animals isoform 1 is expressed  
 CC mainly in the nervous system.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF051335; AAF01564.1; -.  
DR EMBL; AJ242961; CAB71027.1; -.  
DR EMBL; AJ242962; CAB71028.1; -.  
DR EMBL; AJ242963; CAB71029.1; -.  
DR EMBL; AF132045; AAD31019.1; -.  
DR EMBL; AF132046; AAD31020.1; -.  
DR GO; GO:0030176; C:endoplasmic reticulum membrane, intrinsic p. . .; IDA.  
DR GO; GO:0005635; C:nuclear membrane; ISS.  
DR GO; GO:0005515; F:protein binding activity; ISS.  
DR GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.  
DR GO; GO:0030517; P:negative regulation of axon extension; ISS.  
DR InterPro; IPR003388; Reticulon.  
DR Pfam; PF02453; Reticulon; 1.  
DR PROSITE; PS50845; RETICULON; 1.  
KW Endoplasmic reticulum; Alternative splicing; Transmembrane.  
FT DOMAIN 1 989 CYTOPLASMIC (Potential).  
FT TRANSMEM 990 1010 POTENTIAL.  
FT DOMAIN 1011 1104 LUMENAL (Potential).  
FT TRANSMEM 1105 1125 POTENTIAL.  
FT DOMAIN 1126 1163 CYTOPLASMIC (Potential).  
FT DOMAIN 976 1163 RETICULON.  
FT DOMAIN 33 46 POLY-GLU.  
FT DOMAIN 73 76 POLY-ALA.  
FT DOMAIN 140 145 POLY-PRO.  
FT VARSPLIC 1 964 Missing (in isoform 3).  
FT /FTid=VSP\_005656.  
FT VARSPLIC 965 975 AVLSAELSKTS -> MDGQKKHWKDK (in isoform  
FT 3).  
FT /FTid=VSP\_005657.  
FT VARSPLIC 173 975 Missing (in isoform 2).  
FT /FTid=VSP\_005658.  
FT VARSPLIC 192 975 Missing (in isoform 4).  
FT /FTid=VSP\_005659.  
FT CONFLICT 1130 1131 MISSING (IN REF. 3; AAD31020).  
SQ SEQUENCE 1163 AA; 126386 MW; 8CB894B09E94F0B6 CRC64;

Query Match 72.5%; Score 4296.5; DB 1; Length 1163;  
Best Local Similarity 74.0%; Pred. No. 4.2e-158;  
Matches 885; Conservative 104; Mismatches 156; Indels 51; Gaps 19;

QY 1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58  
|||:|||| ||||| ||||| ||||| |||||:||||:|||||  
Db 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEEDDEEEDEDEDLEELEVLERK 60  
QY 59 PAAGLSAAPVPTAPAAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVP 118  
||||||| || ||| |||:||:| ||||| ||||| |||||: ||:|  
Db 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
QY 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPPSTPAAPKR 178  
||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
QY 179 RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSL 224  
||| : |||| |||||:|:||||| ||||| |||||  
Db 167 RGSGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL 226  
QY 225 SPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRLTEFSELEYS 284



Db	227	SPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYS	286
Qy	285	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTKLVKEDE	344
Db	287	EMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KEDR	340
Qy	345	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE	403
Db	341	VVSPEKTMDFNEMQMSVVPVREEYADFKPFEQAWEVKDTYEGSRDVLAA-----RANVE	396
Qy	404	SKVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIAT	463
Db	397	SKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA	455
Qy	464	NIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLTK	523
Db	456	NTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFVLVAVQDSEADYVTTDTLSK	514
Qy	524	VTEEVVANMPEGLTPDLVQACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLC	583
Db	515	VTEAAVSNMPEGLTPDLVQACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLC	574
Qy	584	PSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPENP	642
Db	575	PSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENP	634
Qy	643	PPYEEAMSVSLKVSIGKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDFSD	702
Db	635	PPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSN	694
Qy	703	YSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMVLVKESLTETSFE	762
Db	695	YSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS-E	753
Qy	763	SMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELST	822
Db	754	TVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEFNT	811
Qy	823	AVYSNDDLFIKSEAQIRETETFSDDSPIEIIDFPPTLISSKTDSFSKLAREYTDLEVSHK	882
Db	812	AIYSNDDLSSKEDKIKESETFSDDSPIEIIDFPPTFVSAKDDS-PKLAKEYTDLEVSDK	870
Qy	883	SEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVS	942
Db	871	SEIANIQSGADSLPCLLEPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNVSA	928
Qy	943	LGHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKK	1002
Db	929	L-EPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKK	987
Qy	1003	TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1062
Db	988	TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAY	1047
Qy	1063	LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFN	1122



```

CC      transmembrane domains (By similarity).
CC      -!- ALTERNATIVE PRODUCTS:
CC          Event=Alternative splicing; Named isoforms=1;
CC          Comment=A number of isoforms may be produced;
CC          Name=1;
CC          IsoId=Q99P72-1; Sequence=Displayed;
CC      -!- SIMILARITY: Contains 1 reticulon domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF326337; AAK08076.1; -.
DR      EMBL; AK003859; -; NOT_ANNOTATED_CDS.
DR      MGD; MGI:1915835; Rtn4.
DR      GO; GO:0030176; C:endoplasmic reticulum membrane, intrinsic p. . .; ISS.
DR      GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR      GO; GO:0005635; C:nuclear membrane; ISS.
DR      GO; GO:0005515; F:protein binding activity; ISS.
DR      GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR      GO; GO:0030517; P:negative regulation of axon extension; ISS.
DR      InterPro; IPR003388; Reticulon.
DR      Pfam; PF02453; Reticulon; 1.
DR      PROSITE; PS50845; RETICULON; 1.
KW      Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT      DOMAIN      1      25      CYTOPLASMIC (Potential).
FT      TRANSMEM      26      50      POTENTIAL.
FT      DOMAIN      51      137      LUMENAL (Potential).
FT      TRANSMEM      138      162      POTENTIAL.
FT      DOMAIN      163      199      CYTOPLASMIC (Potential).
FT      DOMAIN      12      199      RETICULON.
SQ      SEQUENCE      199 AA;  22466 MW;  07BE5D580059ED9C CRC64;

```

```

Query Match      15.4%;  Score 915;  DB 1;  Length 199;
Best Local Similarity  97.4%;  Pred. No. 3.3e-29;
Matches 186;  Conservative  2;  Mismatches  3;  Indels  0;  Gaps  0;

```

```

QY      988 KTSVVDLLIYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV I 1047
      |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV I 68

QY      1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLVDLKF 1107
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLVDLKF 128

QY      1108 AVLMMWFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167
      ||||||||||||||||||||||||||||:||||||||||||||||||:|||||||
Db      129 AVLMMWFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQ 188

QY      1168 AKIPGLKRKAE 1178
      |||||||||
Db      189 AKIPGLKRKAE 199

```

RESULT 4

RTN1\_RAT

ID RTN1\_RAT STANDARD; PRT; 777 AA.  
AC Q64548; Q64547;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Reticulon 1 (Neuroendocrine-specific protein) (S-rex).  
GN RTN1 OR NSP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).  
RC STRAIN=Wistar; TISSUE=Brain cortex;  
RX MEDLINE=96386034; PubMed=8793864;  
RA Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,  
RA Georgiev G.P., Buchman V.L.;  
RT "Intracellular compartmentalization of two differentially spliced s-  
RT rex/NSP mRNAs in neurons.";  
RL Mol. Cell. Neurosci. 7:289-303(1996).  
CC -!- FUNCTION: MAY BE INVOLVED IN NEUROENDOCRINE SECRETION OR IN  
CC MEMBRANE TRAFFICKING IN NEUROENDOCRINE CELLS.  
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE (BY  
CC SIMILARITY).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=RTN1-B; Synonyms=S-RexB;  
CC IsoId=Q64548-1; Sequence=Displayed;  
CC Name=RTN1-S; Synonyms=S-RexS;  
CC IsoId=Q64548-2; Sequence=VSP\_005647, VSP\_005648;  
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND  
CC PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS  
CC HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.  
CC EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL  
CC TYPES.  
CC -!- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE  
CC HINDBRAIN AND IN E11 IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC  
CC DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN  
CC THE HINDBRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB  
CC DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE  
CC HINDBRAIN.  
CC -!- SIMILARITY: Contains 1 reticulon domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U17604; AAC53046.1; -.  
DR EMBL; U17603; AAC53045.1; -.  
DR InterPro; IPR003388; Reticulon.

DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane.  
 FT TRANSMEM 604 624 POTENTIAL.  
 FT TRANSMEM 727 747 POTENTIAL.  
 FT DOMAIN 590 777 RETICULON.  
 FT DOMAIN 610 613 POLY-LEU.  
 FT VARSPLIC 1 569 Missing (in isoform RTN1-S).  
 FT /FTId=VSP\_005647.  
 FT VARSPLIC 570 589 IPGPLGSDLVPPLPFFNKQK -> MQATADSTKMDCVWSNW  
 FT KSQ (in isoform RTN1-S).  
 FT /FTId=VSP\_005648.  
 SQ SEQUENCE 777 AA; 83001 MW; AF7479C50F28D0AC CRC64;

Query Match 13.4%; Score 791; DB 1; Length 777;  
 Best Local Similarity 26.6%; Pred. No. 1e-23;  
 Matches 290; Conservative 129; Mismatches 308; Indels 364; Gaps 39;

QY 134 PEDDEPPARPPPPPPASVSPQAE-PVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKI 192  
 | : : | | : | | | | | : | |  
 Db 5 PDLQDEPLSPANPGSQLFGGRGEGEEATPKGARPAQQDGEPAWGS--GAGAGVVS----- 57

QY 193 MDLKEQPGNTISAGQEDFPSVLLETA----ASXP-SLSPLSAASFKEHE---YLGNLSTV 244  
 : : | | : | | | : | : | : | : | :  
 Db 58 -----SRGLCSGPARSPVAMETASTGVAAPDALDHSSSPTLKDGEACYSLSIDI 110

QY 245 L--PTE-----GTLQE-----NVSEASKEVSEKAKTLLIDRLTEFSELEYSEMGSFS 291  
 | | | | : | : : | : | : | | | : | :  
 Db 111 CYPRED SAYFTGILQKENGHITTSSEPEELGTPGPS-LPEVPGTEPHGLLSSDSGIEMT 169

QY 292 VSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLKVEDEVVSSEKA 351  
 : | : : : : : : : : : : : : : : : :  
 Db 170 PAESTEVNKILADPLDQ----- 186

QY 352 KDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVDKKCF 411  
 | : | : | | : | | : | | |  
 Db 187 -----MKAFAECKYIDITRPQEAQGQEEQSPGL----- 213

QY 412 ADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATES--IATNIFPLL 469  
 : | : : : : : | | : | | : | : | :  
 Db 214 -----EDKDLDFKDKDSEVSTKPEGVH-----APNQPSPVEGKLIKDNLF--- 253

QY 470 EDPTSENXTDEKKIEEKKAIQVTE--KNTSTKTSNPPFFVAAQDSETDYVTTDNLTKVTEE 527  
 | : | | | : : : | | | : | | : :  
 Db 254 EESTFAPYIDELSDEQHRMSLVTA PVKITLTEIGPPVMTATHET----- 297

QY 528 VVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFE 587  
 : | | | | | | | | | : | : | :  
 Db 298 ----IPE-----KQDL-----CLKPSPDTPVPTVT 317

QY 588 ESEA---TPSPVLPDIVMEAPLNSAVPSAGASVIQPS SSP-----LEASSVNYESI KH 637  
 | | : | | | : | | | : | | | : | : | :  
 Db 318 VSEPEDDSPG SVTP-----PSSGTEPSAAESQKG SVSEDELIAAIKEAKGLSYET--- 368

QY 638 EPENPPPYEEAMSVSLKVSIGIKEEIKE-----PENINAALQETEAPYISIACDLIKET 690  
 | : | | : | : : | : : | : | : | : | :  
 Db 369 -TESPRPVGQAAD-----RPKVKARSGLP TIPSLLDQEASSAESGDSEI--ELVSED 417

Qy 691 KLSAEPAPDFSDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSI PDVPQKQDETVM 750  
 Db 418 PMASEDALP-SGYVSFGHVSGPPP-----SPAS-----PSIQYSILREEREAE 459

Qy 751 LVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKK 810  
 Db 460 LDSELIIESCDASSASEESPKREQDSPPM--KPGV-----LD-----AIREETSSRATE 506

Qy 811 EKIPLQMEELSTAVYSNDDLFI SKEAQIRETETFS DSSPIE---IIDEFPTLISSKTDSF 867  
 Db 507 ERAP-----SHQGPVEPDPILSFTPTVLQSRPE-- 534

Qy 868 SKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDL SLKNIQPKVEEKISFSDDFSKNG 927  
 Db 535 -----PSSGDGAPVPEPP-----KSQQQKPEEEAVSS---SQSP 565

Qy 928 SATSKVLLLPPDV SALGHTQAEIESIVKPKVLEKEAEKKLP SDTEKEDRSPSAIFSADLG 987  
 Db 566 AAT-----EIPGPLGSDLVPPLPFFN---- 586

Qy 988 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV I 1047  
 Db 587 KQKAIDLLYWRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFRIYKSVL 646

Qy 1048 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDL SLKF 1107  
 Db 647 QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQLYVNSTLKE LRRLFLVQDLVDL SLKF 706

Qy 1108 AVLMMWFTTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQ 1167  
 Db 707 AVLMWLLTYVGALFNGLTLLLMAVVSMTLPVVVYVKHQAQVDQYLGLVRTHINTVVAKI Q 766

Qy 1168 AKIPGLKRKAE 1178  
 Db 767 AKIPGAKRHAE 777

# RESULT 5

## RTN1\_HUMAN

ID RTN1\_HUMAN; STANDARD; PRT; 776 AA.  
 AC Q16799; Q16800; Q16801;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Reticulon 1 (Neuroendocrine-specific protein).  
 GN RTN1 OR NSP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS RTN1-A; RTN1-B AND RTN1-C).  
 RC TISSUE=Lung carcinoma;  
 RX MEDLINE=93293865; PubMed=7685762;  
 RA Roebroek A.J.M., Van de Velde H.J.K., Van Bokhoven A., Broers J.L.V.,  
 RA Ramaekers F.C.S., Van de Ven W.J.M.;

RT "Cloning and expression of alternative transcripts of a novel  
RT neuroendocrine-specific gene and identification of its 135-kDa  
RT translational product.";  
RL J. Biol. Chem. 268:13439-13447(1993).  
RN [2]  
RP ALTERNATIVE SPLICING.  
RX MEDLINE=96429995; PubMed=8833145;  
RA Roebroek A.J.M., Ayoubi T.A.Y., Van de Velde H.J.K.,  
RA Schoenmakers E.F.P.M., Pauli I.G.L., Van de Ven W.J.M.;  
RT "Genomic organization of the human NSP gene, prototype of a novel gene  
RT family encoding reticulons.";  
RL Genomics 32:191-199(1996).  
RN [3]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=98228245; PubMed=9560466;  
RA Hens J., Nuydens R., Geerts H., Senden N.H., Van de Ven W.J.M.,  
RA Roebroek A.J., van de Velde H.J., Ramaekers F.C., Broers J.L.;  
RT "Neuronal differentiation is accompanied by NSP-C expression.";  
RL Cell Tissue Res. 292:229-237(1998).  
CC -!- FUNCTION: MAY BE INVOLVED IN NEUROENDOCRINE SECRETION OR IN  
CC MEMBRANE TRAFFICKING IN NEUROENDOCRINE CELLS.  
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=RTN1-A; Synonyms=NSP-A;  
CC IsoId=Q16799-1; Sequence=Displayed;  
CC Name=RTN1-B; Synonyms=NSP-B;  
CC IsoId=Q16799-2; Sequence=VSP\_005644;  
CC Name=RTN1-C; Synonyms=NSP-C;  
CC IsoId=Q16799-3; Sequence=VSP\_005645, VSP\_005646;  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL AND NEUROENDOCRINE TISSUES  
CC AND CELL CULTURES DERIVED THEREFROM. EXPRESSION OF ISOFORM RTN1-C  
CC IS STRONGLY CORRELATED WITH NEURONAL DIFFERENTIATION.  
CC -!- PTM: ISOFORMS RTN1-A AND RTN1-B ARE PHOSPHORYLATED.  
CC -!- SIMILARITY: Contains 1 reticulon domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L10333; AAA59950.1; -.  
DR EMBL; L10334; AAA59951.1; -.  
DR EMBL; L10335; AAA59952.1; -.  
DR PIR; A46583; A46583.  
DR PIR; I60904; I60904.  
DR Genew; HGNC:10467; RTN1.  
DR MIM; 600865; -.  
DR GO; GO:0030176; C:endoplasmic reticulum membrane, intrinsic p. . .; TAS.  
DR GO; GO:0004871; F:signal transducer activity; NAS.  
DR GO; GO:0030182; P:neuron differentiation; TAS.  
DR GO; GO:0007165; P:signal transduction; NAS.  
DR InterPro; IPR003388; Reticulon.  
DR Pfam; PF02453; Reticulon; 1.

DR PROSITE; PS50845; RETICULON; 1.  
KW Endoplasmic reticulum; Alternative splicing; Transmembrane;  
KW Phosphorylation.  
FT TRANSMEM 603 623 POTENTIAL.  
FT TRANSMEM 726 746 POTENTIAL.  
FT DOMAIN 589 776 RETICULON.  
FT DOMAIN 609 612 POLY-LEU.  
FT VARSPLIC 1 420 Missing (in isoform RTN1-B).  
FT /FTId=VSP\_005644.  
FT VARSPLIC 1 568 Missing (in isoform RTN1-C).  
FT /FTId=VSP\_005645.  
FT VARSPLIC 569 588 GPGPLGPGAPPPLLFLNKQK -> MQATADSTKMDCVWSNW  
FT KSQ (in isoform RTN1-C).  
FT /FTId=VSP\_005646.  
SQ SEQUENCE 776 AA; 83617 MW; CA5B6232353096FE CRC64;

Query Match 13.3%; Score 787.5; DB 1; Length 776;  
Best Local Similarity 31.9%; Pred. No. 1.4e-23;  
Matches 229; Conservative 91; Mismatches 190; Indels 207; Gaps 25;

Qy 588 ESEATPSPVLPDI--VMEAPLNSA-----VPSAGASVIQSSSPLE---ASSVNY---- 632  
| | | | | : | : | : | : | : |  
Db 141 EELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKLADPLDQMKAEAYKYIDIT 200  
Qy 633 --ESIKHEPENPPPYEEA-----MSVSLKVSGIKEEIKEPENINAAL-----QET 675  
| : | | : : | : | : | | : : | : |  
Db 201 RPEEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDK-PAPVEGKI IKDHILLEEST 259  
Qy 676 EAPYISIIACDL-----IKETKLSAEPAPDFSDYSEMAKVEQVPDPHSELVED 722  
| | | | | | : | | | : : : : | : :  
Db 260 FAPYID---DLSEEQRRAPIITTPVKITLLEIPEVSVE-----TTTQEKTPKQDICK 309  
Qy 723 SSPDSEPVDLFS---DDS-----IPDVPQKQ-----DETVMLVKESLTETSF 761  
| | | : | | | | : | | : : | :  
Db 310 PSPDTPVTVTVSEPEDDSPGSITPPSSGTEPSAAESQKGKSI SEDELITAIKEA----- 363  
Qy 762 ESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELS 821  
: | | | : : | : | | | : : :  
Db 364 -KGLSYETAENPRPVGQLADRP-----EVKARSGPPTIPSPLDHEA 403  
Qy 822 TAVYSNDDLFI SKEAQIRETETFS DSSPIEI IDEFP---TLISSKTDSFS----- 868  
: : | | | | : | | : | |  
Db 404 SSAESGD-----SEIELVSEDPMMAEDALPSGYVSFGHVGGPPPPSP 444  
Qy 869 -----KLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSD 921  
: | | : | : : | : : | : | | | | |  
Db 445 ASPSIQYSILREEREAE L DSELI I E SCDASSAS-----EESPKREQDSPPMKPSALD 496  
Qy 922 DF-----SKNGSATS K VLL-----LPPDV SALGHTQAEIESIVKP 956  
| : | | | | | | | | | : | : |  
Db 497 AIREETGVRAEERAPSRRGLAEPGSFLDYPSTEPQPGPELPPGDGAL-----EPETPMLP 551  
Qy 957 KVLEKEAEKKLPSDTEKEDRSPSA-----IFSADLGKTSVVDLLYWRDIK 1001  
: | : : : | : | : | | | | |  
Db 552 -----RKPEEDSSSNQSPAATKGPGPLGPGAPPPLLF---LNKQKAIDLLYWRDIK 599  
Qy 1002 KTG VVFGASL F L L L S L T V F S I V S V T A Y I A L L S V T I S F R I Y K G V I Q A I Q K S D E G H P F R A 1061



```

      :||:||||: | || ||| ||:|||| ||:|||| || ||||| ||| ||:||||:|||||:|
Db      600 QTGIVFGSFLLLLFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKA 659

Qy      1062 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1121
      ||| |: |:|:| :|||:: :|| |:||||| ||| ||||| |||||: |||||
Db      660 YLELEITLSQEIQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKFAVLMWLLTYVGALF 719

Qy      1122 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178
      |||||:::|::|::|::|:| :||||| |||| :| :||| ||||| || ||
Db      720 NGLTLLLMAVVSMFTLPVVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHAE 776

```

# RESULT 6

## RTN3\_MOUSE

```

ID   RTN3_MOUSE          STANDARD;          PRT;    237 AA.
AC   Q9ES97;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Reticulon protein 3.
GN   RTN3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Huang X., Zhou Y., Qiang H., Yuan J., Qiang B.;
RT   "Cloning and expression profile of a novel mouse cDNA encoding a human
RT   RTN3 homolog.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AF195940; AAG31360.1; -.
DR   MGD; MGI:1339970; Rtn3.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Transmembrane; Endoplasmic reticulum.
FT   TRANSMEM           69      89      POTENTIAL.
FT   TRANSMEM          167     187      POTENTIAL.
FT   DOMAIN             49     237      RETICULON.
SQ   SEQUENCE    237 AA;  25428 MW;  EB60A0A7AC45F0DE CRC64;

```

```

Query Match          10.6%;  Score 626;  DB 1;  Length 237;
Best Local Similarity  53.2%;  Pred. No. 5.1e-18;
Matches 125;  Conservative  41;  Mismatches  53;  Indels  16;  Gaps  4;

```

Qy 960 EKEAEKKLPS-DTEKEDRSPSAIFSAD-----LGKTS-----VVDLLYWRDIKKTG 1004  
 | | : | | : | | : | | | | : | | : | | : | | :  
 Db 3 ESSAATQSPSVSSSSSSGAEPSALGGGGSPGACPALGAKSCGSSCAVHDLIFWRDVKKTG 62

Qy 1005 VVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLE 1064  
 | | : | : | | | | : | : | | | | | | : | | : | | : | | : | | :  
 Db 63 FVFGTTLIMLLSLAAFSVISVSYLILALLSVTISFRVYKSVIQAVQKSEEGHPFKAYLD 122

Qy 1065 SEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGL 1124  
 : : : | | | : | : | | : | : | | | : | | | | | | : | | : | | : | | :  
 Db 123 VDITLSSEAFHNYMNAAMVHVNKALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFNGI 182

Qy 1125 TLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-KRKAE 1178  
 | | | | : : | | : | : | : | | | : | : | : | | : | : | : | : | :  
 Db 183 TLLILAE LLVFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKAE 237

# RESULT 7

## RTN3\_HUMAN

ID RTN3\_HUMAN STANDARD; PRT; 236 AA.  
 AC O95197;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Reticulon protein 3 (Neuroendocrine-specific protein-like 2) (NSP-like  
 DE protein II) (NSPLII).  
 GN RTN3 OR NSPL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Retina;  
 RX MEDLINE=99265974; PubMed=10331947;  
 RA Moreira E.F., Jaworski C.J., Rodriguez I.R.;  
 RT "Cloning of a novel member of the reticulon gene family (RTN3): gene  
 RT structure and chromosomal localization to 11q13.";  
 RL Genomics 58:73-81(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Huang X., Zhou Y., Du G., Yuan J., Qiang B.;  
 RT "Cloning and expression analysis of a cDNA encoding a novel  
 RT neuroendocrine-specific protein-like protein 1: NSPL1.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, Eye, and Lymph;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum (Potential).  
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN  
 CC BRAIN. THREE TIMES MORE ABUNDANT IN MACULA THAN IN PERIPHERAL  
 CC RETINA.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF059524; AAC99319.1; -.  
 DR EMBL; AF059529; AAD20951.1; -.  
 DR EMBL; AF059525; AAD20951.1; JOINED.  
 DR EMBL; AF059526; AAD20951.1; JOINED.  
 DR EMBL; AF059527; AAD20951.1; JOINED.  
 DR EMBL; AF059528; AAD20951.1; JOINED.  
 DR EMBL; AF119297; AAD26810.1; -.  
 DR EMBL; BC000634; AAH00634.1; -.  
 DR EMBL; BC010556; AAH10556.1; -.  
 DR EMBL; BC011394; AAH11394.1; -.  
 DR EMBL; BC022993; AAH22993.1; -.  
 DR Genew; HGNC:10469; RTN3.  
 DR MIM; 604249; -.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Transmembrane; Endoplasmic reticulum.  
 FT TRANSMEM 68 88 POTENTIAL.  
 FT TRANSMEM 177 197 POTENTIAL.  
 FT DOMAIN 48 236 RETICULON.  
 SQ SEQUENCE 236 AA; 25609 MW; DDC6A4544ABCD7B7 CRC64;

Query Match 10.6%; Score 625.5; DB 1; Length 236;  
 Best Local Similarity 54.8%; Pred. No. 5.4e-18;  
 Matches 119; Conservative 41; Mismatches 56; Indels 1; Gaps 1;

QY 963 AEKKLPSTDEKEDRSPSAIFSDALGKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSI 1022  
 || | : : :| ||::|||:|||| ||| :| :|||| ||:

Db	20	AEPSAPGGGGSPGACPALGKSCSSSCAVHDLFIWRDVKKTGTFVGGTTLIMLLSLAAFSV	79
Qy	1023	VSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSAL	1082
		:   : :                   :    :    :      : :	
Db	80	ISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAM	139
Qy	1083	GHVNCTIKELRRLFLVDDLVDLSLKFAVLMMWFVTFYVGALFNGLTLLILALISLFSVPVIYE	1142
		:  :  :       :         :       :    :      : :    :	
Db	140	VHINRALKLIIRLFLVEDLVDLSKLAVFMWLMTYVGAVFNGITLLILAELLIFSVPPIVYE	199
Qy	1143	RHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-KRKA	1178
		:::       : :    :       :    : :	
Db	200	KYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKAE	236

## RESULT 8

RTN2 HUMAN

```

ID      RTN2_HUMAN          STANDARD;          PRT;      545 AA.
AC      O75298; O60509;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE      protein 1) (NSPLI).
GN      RTN2 OR NSPL1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC      TISSUE=Lung carcinoma;
RX      MEDLINE=98360096; PubMed=9693037;
RA      Roebroek A.J.M., Contreras B., Pauli I.G.L., Van de Ven W.J.M.;
RT      "cDNA cloning, genomic organization, and expression of the human RTN2
RT      gene, a member of a gene family encoding reticulons.";
RL      Genomics 51:98-106(1998).
RN      [2]
RP      SEQUENCE OF 108-545 FROM N.A. (ISOFORM RTN2-B).
RC      TISSUE=Brain;
RX      MEDLINE=98191726; PubMed=9530622;
RA      Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT      "Molecular cloning of a novel mouse gene with predominant muscle and
RT      neural expression.";
RL      Mamm. Genome 9:274-282(1998).
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC      reticulum (Potential).
CC      -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=RTN2-A;
CC      IsoId=O75298-1; Sequence=Displayed;
CC      Note=Isoform RTN2-C is produced by alternative initiation at
CC      Met-341 of isoform RTN2-A;
CC      Name=RTN2-B;
CC      IsoId=O75298-2; Sequence=VSP_005649;
CC      Event=Alternative initiation;
CC      Comment=2 isoforms, RTN2-A (shown here) and RTN2-C, are produced

```

CC by alternative initiation at Met-1 and Met-341;  
 CC -!- TISSUE SPECIFICITY: ISOFORM RTN2-C IS HIGHLY EXPRESSED IN SKELETAL  
 CC MUSCLE.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF004222; AAC32542.1; -.  
 DR EMBL; AF004223; AAC32543.1; -.  
 DR EMBL; AF004224; AAC32544.1; -.  
 DR EMBL; AF038540; AAC14910.1; -.  
 DR Genew; HGNC:10468; RTN2.  
 DR MIM; 603183; -.  
 DR GO; GO:0030176; C:endoplasmic reticulum membrane, intrinsic p. . .; NAS.  
 DR GO; GO:0004871; F:signal transducer activity; NAS.  
 DR GO; GO:0007165; P:signal transduction; NAS.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane;  
 KW Alternative initiation.  
 FT CHAIN 1 545 RETICULON PROTEIN 2, ISOFORM RTN2-A.  
 FT CHAIN 341 545 RETICULON PROTEIN 2, ISOFORM RTN2-C.  
 FT INIT\_MET 341 341 FOR ISOFORM RTN2-C.  
 FT TRANSMEM 368 388 POTENTIAL.  
 FT TRANSMEM 463 483 POTENTIAL.  
 FT DOMAIN 345 545 RETICULON.  
 FT VARSPLIC 272 344 Missing (in isoform RTN2-B).  
 FT /FTid=VSP\_005649.  
 SQ SEQUENCE 545 AA; 59263 MW; 971FD2F909E1E9E6 CRC64;

Query Match 8.7%; Score 514; DB 1; Length 545;  
 Best Local Similarity 28.2%; Pred. No. 2.9e-13;  
 Matches 169; Conservative 81; Mismatches 186; Indels 164; Gaps 18;

QY 676 EAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPVDHSELVEDSSPD-SEPVDL-F 733  
 ||| | | | | | : | : | : | | : | | : | |  
 Db 13 EAP-----STASSTPDSTEGGNDDSDFRELHTAREFSEEDDEE--ETTSQDWGTPRELTF 64  
 QY 734 S-----DDSI PDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEG 780  
 | | | | : | : | : | : | : | | |  
 Db 65 SYIAFDGVVGSGGRRDSTARRPRPQGRSVSEPRDQHPQPSLGDSLESIPSLSQSPEPGRR 124  
 QY 781 GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFIKSEAQIRE 840  
 | | | | | | : | : : | : : : |  
 Db 125 GDP-----DTAPPSE---RPLEDLRLRLDHLG-----WVARGTGSGE 158  
 QY 841 TETFSDDSSPIEIIDFPTLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTEL 900  
 : | | : | | | : | | | | | |  
 Db 159 DSSTSSSTPLE--DEEP-----QEPNRLETGEAGE-----EL 188

Qy 901 PHDLSLKNIQPKVEEKI-----SFSDDFSKNGSATSQVLLLLPPDVSA LGHT 946  
 || |: || | : : | |: : : : || : |  
 Db 189 --DLRLRLAQPSSEVLTPQLSPGSGTPQAGTPSPSRSDSNSGPPEEPLLEEEEEKQWG PL 246

Qy 947 QAE-----IESIVKPKVLEKEAE-----KKLP----- 968  
 : | : | |: |: | : | : |  
 Db 247 EREPVRGQCLDSTDQLEFTVEPRLLGTAMEWLKTSLLLAVYKTVPIELSPPLWTAIGWV 306

Qy 969 -----SDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVFGASL 1011  
 : : : || |: | : | ||||: | : : |||| :  
 Db 307 QRGPTPTPVLRVLLKWAKSPRSSGVPSSLGADMG-SKVADLLYWKDTRTSGVVFTGLM 365

Qy 1012 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISE 1071  
 || | ||||| |: || || || |: |: |: | : |: |: |: : : : :  
 Db 366 VSLLCLLHFSIVSVA AHLALLLLCGTISLRVYRKVLQAVHRGDGANPFQAYLDVDLTLTR 425

Qy 1072 ELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILAL 1131  
 | : : | : : || ||: ||||| |: | : : |: |: |: ||||| :  
 Db 426 EQTERLSHQITSRVVSAATQLRHFFLVEDLVDLSLKLALLFYILTFVGAI FNGLTLLILGV 485

Qy 1132 ISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-----KRKAE 1178  
 | ||: |: |: | : ||||| |: || : ||: ||||| | |||  
 Db 486 IGLFTIPLLRYRQHQAQIDQYVGLVTNQLSHIKAKIRAKIPGTGALASAAA AVSGSKAKAE 545

# RESULT 9

## RTN2\_MOUSE

ID RTN2\_MOUSE STANDARD; PRT; 471 AA.  
 AC O70622; O70620;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like protein 1) (NSPL1).  
 GN RTN2 OR NSPL1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.  
 RC STRAIN=FVB/N, and 129/Sv; TISSUE=Cerebellum, and Skeletal muscle;  
 RX MEDLINE=98191726; PubMed=9530622;  
 RA Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;  
 RT "Molecular cloning of a novel mouse gene with predominant muscle and  
 RT neural expression."  
 RL Mamm. Genome 9:274-282(1998).  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum  
 CC (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=Brain;  
 CC IsoId=O70622-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Muscle;  
 CC IsoId=O70622-2; Sequence=VSP\_005650, VSP\_005651;  
 CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN NEURAL AND MUSCULAR  
 CC TISSUES.

```

CC  -!- SIMILARITY: Contains 1 reticulon domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF038537; AAC14906.1; -.
DR  EMBL; AF038537; AAC14907.1; -.
DR  EMBL; AF038538; AAC14908.1; -.
DR  EMBL; AF038539; AAC14909.1; -.
DR  EMBL; AF093624; AAD13195.1; -.
DR  MGD; MGI:107612; Rtn2.
DR  InterPro; IPR003388; Reticulon.
DR  Pfam; PF02453; Reticulon; 1.
DR  PROSITE; PS50845; RETICULON; 1.
KW  Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT  TRANSMEM      295      315      POTENTIAL.
FT  DOMAIN        272      471      RETICULON.
FT  VARSPLIC      1        267      Missing (in isoform 2).
FT                                     /FTid=VSP_005650.
FT  VARSPLIC      268      271      PLLL -> MGSK (in isoform 2).
FT                                     /FTid=VSP_005651.
SQ  SEQUENCE      471 AA;  51346 MW;  9BBD8F372CF63AD3 CRC64;

```

```

Query Match          8.4%;  Score 500;  DB 1;  Length 471;
Best Local Similarity 29.3%;  Pred. No. 8.3e-13;
Matches 159;  Conservative 83;  Mismatches 177;  Indels 124;  Gaps 16;

```

```

QY      676 EAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQPVPDHSELVEDSSPD-SEPVDL-F 733
      |||      | | | | | : : : | | | : | | | : | |
Db      13 EAP-----STASSTPDSTEGGNDSDSFRELHTAREFSEDEEE--ETTSQDWGTPRELTF 64

QY      734 S-----DDSDPDVPQKQDETVMVLVKESLTETSFESMIEYENKEKLSALPPEG 780
      |      ||| : | : | : : : : | | |
Db      65 SYIAFDGVVGSGGRRDSVVRPRPQGRSVSEPRDPPQQSGLGDSLESIPSLSQSPPEGR 124

QY      781 GKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFIKSEAQIRE 840
      | |      | : | | | : | : | : | : | : | : | : |
Db      125 GDP-----DPVPPAE---RPLEELRLRLDQLG-----WVVRSAAGSGE 158

QY      841 TETFSDDSSPIEIIIDEFPTLISS----KTDSFSKLARE-YTDLEV-----SHKSEIAN 887
      | | : | : | : | : : : : | : : | : | : | : | : |
Db      159 DSATSSSTPLE--NEEPDGLEASEAGEETNLELRLAQSLHLQLEVLTPQLSPSSGTPQAH 216

QY      888 APDGAGSLPCTELPHDL  QPKVEEKISFSDDFSKNGSAZSKVLLLPDVSALGHTQ 947
      | | | | | | : | | | : | : | : | : | : | : |
Db      217 TPSPQRSQDSNSGPDDEPLLN---VEEH-----WRLLEQEPITAQCLDST 259

QY      948 AEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKKTGVVF 1007
      : | : : | : |      | | | | | : : | |
Db      260 DQSEFMLEPLLL-----VADLLYWKDTRTSGAVF 288

QY      1008 GASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEV 1067

```

```

      : || | ||||| |::| | || |:: |::: | :||:|::: ::
Db      289 TGLMASLLCLLHFSIVSVAHLALLGLCATISLRVYRKVLQAVHRGDGTNPFFQAYLDMDL 348
Qy      1068 AISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLL 1127
      :: | :: | || | :|| |||:|||||| |:: |:: |::|:|||||:
Db      349 TLTREQTERLSQQIASHVVSTATQLRHFFLVEDLVDSLKLALLFYILTFVGAIFNGLTLV 408
Qy      1128 ILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-----KR 1175
      || ::||:|::|: |::| ||| |:: |:: |::|:|||||
Db      409 ILGVVALFTVPLLYRQHQAQIDQYVGLVTNQLSHIKAKIRAKIPGTGTLAPTASVSGSKA 468
Qy      1176 KAE 1178
      |||
Db      469 KAE 471

```

# RESULT 10

## CPN\_DROME

```

ID   CPN_DROME          STANDARD;      PRT;      865 AA.
AC   Q02910;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Calphotin.
GN   CPN OR CAP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Canton-S;
RX   MEDLINE=93165729; PubMed=8094559;
RA   Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
RT   "Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Canton-S;
RX   MEDLINE=93165730; PubMed=8434015;
RA   Ballinger D.G., Xue N., Harshman K.D.;
RT   "A Drosophila photoreceptor cell-specific protein, calphotin, binds
RT   calcium and contains a leucine zipper.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
CC   -!- FUNCTION: Might function as a calcium-sequestering "sponge" to
CC   regulate the amount of free cytoplasmic calcium. It binds 0.3 mole
CC   of Ca(2+) per mole of protein.
CC   -!- SUBUNIT: Homodimer (Probable).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPARTMENT.
CC   -!- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS OF
CC   COMPOUND EYES AND OCELLI.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL
CC   DEVELOPMENT.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L02111; AAA28405.1; -.  
DR EMBL; L05080; AAA28420.1; -.  
DR PIR; A47282; A47282.  
DR PIR; A47283; A47283.  
DR FlyBase; FBgn0010218; Cpn.  
DR GO; GO:0005509; F:calcium ion binding activity; IDA.  
KW Calcium-binding.  
FT CONFLICT 36 36 A -> AVAPAVVA (IN REF. 2).  
FT CONFLICT 43 43 I -> T (IN REF. 2).  
FT CONFLICT 64 64 I -> V (IN REF. 2).  
FT CONFLICT 76 76 T -> A (IN REF. 2).  
FT CONFLICT 100 100 P -> PP (IN REF. 2).  
FT CONFLICT 126 127 VQ -> AP (IN REF. 2).  
FT CONFLICT 154 154 I -> V (IN REF. 2).  
FT CONFLICT 160 160 S -> T (IN REF. 2).  
FT CONFLICT 534 534 A -> E (IN REF. 2).  
FT CONFLICT 699 699 I -> T (IN REF. 2).  
FT CONFLICT 703 703 V -> L (IN REF. 2).  
FT CONFLICT 721 721 D -> E (IN REF. 2).  
SQ SEQUENCE 865 AA; 84781 MW; 2110417E0B0E7CFE CRC64;

Query Match 5.5%; Score 324.5; DB 1; Length 865;  
Best Local Similarity 22.3%; Pred. No. 9.6e-06;  
Matches 229; Conservative 128; Mismatches 375; Indels 297; Gaps 49;

QY 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVP 118  
| : ||| | | ||: | ||| |: | | | |: | | :|  
Db 9 PVSAPVAAPV-TPSAVAAPVQVVSPAAPVAPAAPIAVTPVAPPPTLASVQPATV--TIP 65  
  
QY 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAP----APAAPP--ST 172  
||:|::||:|:| || | |||: || : || | || || :|  
Db 66 APAPIAAASVTP---VASVAPPVVAAPTPPAA-SPVSTPVAVAQIPVAVSAPVAPPVAAT 121  
  
QY 173 PAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSLSPLSAASF 232  
| : : | : | | || | : | || | : || : | :|  
Db 122 PTPVVQIPVAAPVIAT-----PPVAASA-----PT----PAAVTPVISPVIAS-- 160  
  
QY 233 KEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGSSFSV 292  
| : | | | : : | : : | : : | :|  
Db 161 -----PPVVPANTT---VPVAAPVAAVPAAPVVPVAPVLAP-----AV 194  
  
QY 293 SPKAESAVIVAN-----PREEII VKNKDEEEKLVSNILHXQQELPTALTCLKVEDEVVS 347  
: | | : || | || | : | : | : | ||:  
Db 195 APAV--APVVAETPAPPPVAEIPVAT-----IPECVAPLIPEVSVVA 234  
  
QY 348 SEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKVD 407  
: | : | | | | | | : : : : :|  
Db 235 T---KPLAAAEFVVVAPPATET-----PVVAPAAASPHVSVAPAVETAVVAPVS 280  
  
QY 408 KKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFP 467  
|| : : | : || ||

Db 281 -----ASTEPPV---AAATLTAPETPAL----- 301

Qy 468 LLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFVAAQDSETDYVTTDNLTKVTEE 527  
| : | | | |

Db 302 -----APVVAESQ-----VAA-----NTVVATPP 320

Qy 528 VVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPA----- 579  
| | | : | : | : | | : | | | : | |

Db 321 TPAPEPETIAPPVV-----AETPEVASVAVA-ETTPPVV--PPVAAESI-PAPVVATTPV 371

Qy 580 -AQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEASSVNYESIKHE 638  
| | : : | | : : | | : | | : | |

Db 372 PATLAVTDPDVTASAVPELPPVIAAPSPVPSAVAETPVDLAPPVLPVAAEPVPAVVAEET 431

E Qy s639 PENPPPYEAMS SV-SLKVSGIKEEIKEPENINAALQETEAPYISIACDLIKET PA 697  
| | | | : : : | : : | | : | | : | : | |

Db 432 PETPAPASAPVTIAALDIPEVAPVIAAPSDAPAEAPSAAAPIVSTPPTTASVPETTAPPA 491

Qy 698 P-----DFSDYSEMAKVEQVPDHSSEL-----VEDSSPDSEPVDLFSDD SIP--DVPQK 744  
| | | | : | | | : | | : | | : | |

Db 492 AVPTEPIDVSVLSE-AAIETPVAPPVEVTTEVAVADVAPPEAAADLIEPVEPPAPIPDL 550

Qy 745 QDETVMLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEV 804  
: : | : : : | : : : | : : : | : : : |

Db 551 LEQTTSVPAVEAAESTSSPIPE-----TSLPPPNEAVASPEVAVAPITAPEPIPEPEP 603

Qy 805 STLSKKEKIPLQ---MEELSTAVYSNDDLFIKSEAQIRETETFSDDSPIEIIDFPTLI 860  
| : | | : : : | | | : | | | | | | | |

Db 604 SLATPTEPIPVEAPVVIQEAVDV-----EVPVTETST---SIP-ETTVEFPEAV 649

| Qy 861 SSKTDSFSKLA -SHKSEIANAPDGAGSLPCTELP-----:---HDL\$1 906  
: | | | : | : : : | | | | | | : | : :

Db 650 AEKV-----LDPAITEAPVTTQEPDVANINDGA---PATEITTPAVEIVTAAAEVSDIAI 701

Qy 907 KNIQPKVEEKISFSDDFSKNGSATSKVLL---LP-----PDVSALG----- 944  
| | | : : : : : : : : : : | : | |

Db 702 PVIDPPVPQEIAVA-EIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAAEVP 760

Qy 945 -----HTQAEIESIVKPKVLEKEAEKKLPSDTEKEDRSPSAIF SADLGKTSVVDLLY 996  
: | | : | | : | | : | : : | | | : | : | |

Db 761 ITAGDNPNTSVGISEVV-PTIAEKPVEEVPTSEIPEQSSSPSD--SVPVAK--ITPLL- 814

Qy 997 WRDIKKTGV 1005  
| | : : | |

Db 815 -RDLQTTDV 822

# RESULT 11

## PCLO\_HUMAN

ID PCLO\_HUMAN STANDARD; PRT; 5147 AA.  
AC Q9Y6V0; O43373; O60305; Q9BVC8; Q9UIV2; Q9Y6U9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Piccolo protein (Aczonin) (Fragments).  
GN PCLO OR ACZ OR KIAA0559.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 1-759 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99439764; PubMed=10508862;  
 RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,  
 RA Kilimann M.W.;  
 RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active  
 RT zones, shares homology regions with rim and bassoon and binds  
 RT profilin.";  
 RL J. Cell Biol. 147:151-162(1999).  
 RN [2]  
 RP SEQUENCE OF 552-4404 FROM N.A.  
 RA Kraemer J., Wollam C., Wohldmann P., McGrane B.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:31-39(1998).  
 RN [4]  
 RP SEQUENCE OF 4405-4439 FROM N.A.  
 RC TISSUE=Placenta; 75XI a A., Kotani H., TP KD 442850XS  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., AbY ., Mullahy S-59XP2113  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 4405-5147 FROM N.A.  
 RA Kalicki J., Elliott G.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle

```

CC      trafficking (By similarity).
CC      -!- SUBUNIT: Interacts with Rabac1/Pral and profilin (By similarity).
CC      -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC      synaptic junctions (By similarity).
CC      -!- ALTERNATIVE PRODUCTS:
CC          Event=Alternative splicing; Named isoforms=2;
CC          Comment=Additional isoforms seem to exist;
CC          Name=1;
CC              IsoId=Q9Y6V0-1; Sequence=Displayed;
CC          Name=2;
CC              IsoId=Q9Y6V0-2; Sequence=VSP_003923, VSP_003924, VSP_003925,
CC                  VSP_003926, VSP_003927;
CC          Note=No experimental confirmation available;
CC      -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC      phospholipids. Calcium binds with low affinity but with high
CC      specificity and induces a large conformational change.
CC      -!- SIMILARITY: C                C2 domains.                                KNY 2003
CC      -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y19188; CAB60727.1; -.
DR      EMBL; AC004903; AAD20936.1; -.
DR      EMBL; AC004886; AAD21789.1; -.
DR      EMBL; AB011131; BAA25485.1; -.
DR      EMBL; BC001304; AAH01304.1; -.
DR      EMBL; AC004082; AAB97937.1; -.
DR      PIR; T00634; T00634.
DR      HSSP; P04410; 1A25.
DR      Genew; HGNC:13406; PCLO.
DR      MIM; 604918; -.
DR      GO; GO:0005856; C:cytoskeleton; NAS.
DR      GO; GO:0045202; C:synaptic junction; ISS.
DR      GO; GO:0005509; F:calcium ion binding activity; ISS.
DR      GO; GO:0005544; F:calcium-dependent phospholipid binding acti. . .; ISS.
DR      GO; GO:0005522; F:profilin binding activity; ISS.
DR      GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
DR      GO; GO:0016080; P:synaptic vesicle targeting; ISS.
DR      InterPro; IPR000008; C2.
DR      InterPro; IPR001565; Synaptotagmin.
DR      PRINTS; PR00360; C2DOMAIN.
DR      PRINTS; PR00399; SYNAPTOTAGMN.
DR      SMART; SM00239; C2; 2.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS50004; C2_DOMAIN_2; 2.
KW      Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
KW      Repeat; Alternative splicing.
FT      NON_TER      1      1
FT      DOMAIN      400      465      10 X 10 AA TANDEM APPROXIMATE REPEATS OF
FT      P-A-K-P-Q-P-Q-Q-P-X.
FT      ZN_FING      499      523      C4-TYPE (POTENTIAL).

```

FT	ZN_FING	969	992	C4-TYPE (POTENTIAL).
FT	NON_CONS	1010	1011	
FT	DOMAIN	2300	2325	POLY-PRO.
FT	DOMAIN	4391	4442	PDZ.
FT	DOMAIN	4544	4633	C2 DOMAIN 1.
FT	DOMAIN	5031	5121	C2 DOMAIN 2.
FT	VARSPPLIC	4404	4404	S -> SGNGLGIRIVGGKEIPGHSGEIGAYIAKILPGGSAE
FT				QTGKLM EG (in isoform 2).
FT				/FTId=VSP_003923.
FT	VARSPPLIC	4534	4534	K -> KPTDGTKVVSHPITGEIQ (in isoform 2).
FT				/FTId=VSP_003924.
FT	VARSPPLIC	4576	4576	G -> GQVMVVQNAS (in isoform 2).
FT				/FTId=VSP_003925.
FT	VARSPPLIC	4757	4761	TAHKS -> SKRRK (in isoform 2).
FT				/FTId=VSP_003926.
FT	VARSPPLIC	4762	5147	Missing (in isoform 2).
FT				/FTId=VSP_003927.
SQ	SEQUENCE	5147 AA;	563537 MW;	CD5D84990498CD3C CRC64;

Query Match 5.4%; Score 320; DB 1; Length 5147;  
 Best Local Similarity 21.8%; Pred. No. 0.00013;  
 Matches 260; Conservative 152; Mismatches 440; Indels 340; Gaps 61;

Qy	8	PLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPAAGLSAAP	67
		:         :   :	
Db	260	PSLPSPSKPPIQQPTPGKPPAQPGHEKSQPG-----PAKPPAQPSGLT	303
Qy	68	VPTAPAAGA---PLMDFGNDFVPPAPRGPLPAAPPV-----	100
		:	
Db	304	KPLAQPGTVKPPVQPPGTTKPPAQPLG--PAKPPAQQTGSEKPSSEQPGPKALAQPPGV	361
Qy	101	--APERQPSWDPSVSSVTPAPSPLS-----AAAVSPSKLPEDDEPPA-----RPP	144
		:     :       :       :     :	
Db	362	GKTPAQQPG-PAKPPTQQVGTGPKPLAQQGLQSPAKAPGPTKTPAQTKPPSQQPGSTKPP	420
Qy	145	P--PPPASVSPQAEPVWTPPAPAP-AAPPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGN	201
		:   :     :	
Db	421	PQQPGPAKPSPPQPGSTKPPSQQPGSAKPSA-----QQPSP	456
Qy	202	TISAGQEDFPSVLLETAASXPSLSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKE	261
		:   :   :   :   :   :   :   :	
Db	457	AKPSAQQ-FTKPVSTGTGFGKPLQPPTVSPSAKQPPSQGLPKTICPL-----CNTTELLH	510
Qy	262	VSEKA-----KTL LI-----DRDLTEFSE-----LEYSEMSSFSVSPKA----	296
		:   : :       :   :	
Db	511	VPEKANFNTCTECQTTVCSLCGFNPNPHLTEAKEWLCLNCQMKRALGGDLAPVPSSPQPK	570
Qy	297	-----ESAVIVANPREEIIVKNKDDEEEKLVSNNILHXQQEL-----PTALTCLKVKED	343
		:   : :       :	
Db	571	LKTAPVTTTSAVSKSSPQQQTSPKKDAAPK-----QDLSKAPEPKKPPPLVKQP	620
Qy	344	EVVSSEKAK-----DSFNEKRVAVEAPMREE-----YADFKPFERVWEVKDSKEDSDM	391
		:       :   :   :       :     :	
Db	621	TLHGSPSAKAKQPPEADSLSKPAPPKEPSVPSEQDKAPVADDKP----KQPKMVKPTTDL	676
Qy	392	LAAGGKIESNLESKVD---KKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYI	448
		: : :   : :   : :   : :   : :	

Db 677 VSS-----SSATTKPDIPSSKVSQAEEKTTPPLKTD SAKPSQSF PPTGEKV----- 723

Qy 449 TCAPFN-----PAATESIATNIFLLE-----DPTSENXTDEKKIEEKKAQIVTEKN 495  
 ||: ||: |:: | | || :|| | |||

Db 724 --TPFDSKAI PRPASDSKI ISHPGPSSES KGQKQVDPV-----QKKEEPKKAQTKMSPK 775

Qy 496 TSTK-----TSNPFFVAAQDSETDYVTTD-----NLTKVTEEVVANMPEGLT 537  
 | | | | : || :| : |

Db 776 PDAKPM PKGSPTPPGPRPTAGQTVPTPQQSPKQE QSRRFSLNLGSITD---APKSQPTT 832

Qy 538 PDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPA--AQLCPSFEESEATPSP 595  
 | || || : : | : | : | : | : |

Db 833 P---QET-----VTGKLFGFGASI -FSQASNLI STAGQPGPHSQSGPGAPMKQA-PAP 880

Qy 596 VLPDIVMEAPLNSA-VPSAGASVI-----QPSSSPLEASSVNYESI K--HEPENPPPY 645  
 | :| | | | :| | : :| : ||

Db 881 SQPPTSQGPPKSTGQAPPAPAKSI PVKKETKAPAAEKL EPKAEQAPT VKRTETEKKPPPI 940

Qy 646 EEAMSVSLKVSIGIKEEIKEPENINAALQETEAPYISIACDLIK-ETKLSAEPAPDFSDYS 704  
 ::: |:: ||: : :| | | | : : : | : :

Db 941 KDSKSLT-----AEPQKAVLPTKLEKSPKPESTCPLCKTEL NIGSKDPPNFNTCT 990

Qy 705 EMAKVE-----QPVPDHSELVE-----DSSPDSEPVDLFSDDS--- 737  
 | | :| : | | :| : : :|

Db 991 ECKNQVCNLCGFNPTPHLTENCQTQRAISGQLGDIRKMPPAPSGPKASMPVPTESSQK 1050

Qy 738 --IPDVPQ-----KQD-----ETVML--VKESLTETSFESMIEYENKEKLSALPPEGG 781  
 :| || ||: | :| ||| :| : :| :| :

Db 1051 TAVP--PQVKLVKKQE QEVKTEAEKVILEKVKETLSMEKI PPMVTTDQKQEESKLEKDKA 1108

Qy 782 KPYLESFKLSLDNTKDTLLPDEVSTLSKKEKI PLQMEELSTAVYSNDDL FISKEAQIRET 841  
 | | | | :| :| : :| | | :| :| :

Db 1109 SALQE--KKPLPEEK-KLIPEE-EKIRSEEKKPLLEEKKPTP---EDKKLLPEAKTSAP 1160

Qy 842 ETFSD---SSPIEIDE-FPTLISSKTD SFSKLAREYTDLEVSHKSEIANAPDGA-GSLP 896  
 | :| :| : :| | : : :| | ||

Db 1161 EEQKHDLLKSQVQIAEEKLEGRVAPKTVQEGK-----QPQTKMEGLPSGTPQSLP 1210

Qy 897 CTELP HDLSLKNI----QPKVEEKISFSDDFS KNGSATS KVL LPPDV SALGHTQ----- 947  
 : | :| || : :| : :| :| :| :

Db 1211 KED---DKTTKTIKEQPQPPCTAKPDQE KEDDKSDTSSSQPKSPQGLSDTGYSSDGISS 1267

Qy 948 --AEIESIVKPKVLEKEAEKKLPDSTEKEDRSPSAIFSADLGK--TSVVDLL 995  
 || |:: ||: | | : :| ||: :| | :| :| :

Db 1268 SLGEIPSLIPTD--EKDILKGLKKDSFSQESSPS--PSDLAKLESTVLSIL 1315

# RESULT 12

## PCLO\_MOUSE

ID PCLO\_MOUSE STANDARD; PRT; 5038 AA.  
 AC Q9QYX7; Q9QYX6; Q9QZJ0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Piccolo protein (Presynaptic cytomatrix protein) (Aczonin) (Brain-  
 DE derived HLMN protein).  
 GN PCLO OR ACZ.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, ALTERNATIVE SPLICING,  
 RP TISSUE SPECIFICITY, AND INTERACTION WITH PROFILIN.  
 RC TISSUE=Brain;  
 RX MEDLINE=99439764; PubMed=10508862;  
 RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,  
 RA Kilimann M.W.;  
 RT900XKczonin, ~~5155W-kdHpnagtcive sheiffldiHapgofoejnJufGpresynaptic~~ 66  
 RT zones, shares homology regions with rim and bassoon and binds  
 RT profilin.";  
 RL J. Cell Biol. 147:151-162(1999).  
 RN [2]  
 RP REVISIONS.  
 RC TISSUE=Brain;  
 RA Kilimann M.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 4502-4682 FROM N.A.  
 RC TISSUE=Brain;  
  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle  
 CC trafficking.  
 CC -!- SUBUNIT: Interacts with Rabac1/Pral and profilin.  
 CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of  
 CC synaptic junctions.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC E0ent=Alt06pat1995 splicing; Named ibonding=2alcium and  
 CC Name=1;  
 CC IsoId=Q9QYX7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9QYX7-2; Sequence=VSP\_003928, VSP\_003929;  
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Low levels found in  
 CC stomach. Not detected in other tissues analyzed including adrenal  
 CC gland, testis and pancreas.  
 CC -!- DOMAIN: C2 domain 1 is  
 CC phospholipids. Calcium binds with low affinity but with high  
 CC specificity and induces a large conformational change.  
 CC -!- SIMILARITY: Contains 2 C2 domains.  
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----6-----eng-----2;-----  
 DR EMBL; Y19185; CAB60731.2; -.  
 DR EMBL; Y19186; CAB60732.2; -.  
 DR EMBL; AF181269; AAD55786.2; -.

DR HSSP; P04410; 1A25.  
 DR MGD; MGI:1349390; Pclo.  
 DR GO; GO:0045202; C:synaptic junction; IDA.  
 DR GO; GO:0005509; F:calcium ion binding activity; ISS.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding acti. . .; ISS.  
 DR GO; GO:0005522; F:profilin binding activity; IDA.  
 DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; IDA.  
 DR GO; GO:0016080; P:synaptic vesicle targeting; NAS.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR001478; PDZ.  
 DR Pfam; PF00168; C2; 2.  
 DR Pfam; PF00595; PDZ; 1.  
 DR SMART; SM00239; C2; 2.  
 DR SMART; SM00228; PDZ; 1.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 2.  
 DR PROSITE; PS50106; PDZ; 1.  
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;  
 KW Repeat; Alternative splicing.  
 FT DOMAIN 371 470 10 X 10 AA TANDEM APPROXIMATE REPEATS OF  
 FT P-A-K-P-Q-P-Q-Q-P-X.  
 FT ZN\_FING 502 526 C4-TYPE (POTENTIAL).  
 FT ZN\_FING 967 990 C4-TYPE (POTENTIAL).  
 FT DOMAIN 2305 2329 POLY-PRO.  
 FT DOMAIN 4394 4488 PDZ.  
 FT DOMAIN 4607 4705 C2 DOMAIN 1.  
 FT DOMAIN 4922 5012 C2 DOMAIN 2.  
 FT VARSPLIC 4829 4833 TKPTN -> SKRRK (in isoform 2).  
 FT /FTId=VSP\_003928.  
 FT VARSPLIC 4834 5038 Missing (in isoform 2).  
 FT /FTId=VSP\_003929.  
 SQ SEQUENCE 5038 AA; 547600 MW; DADA460CF3B40888 CRC64;

Query Match 5.3%; Score 315; DB 1; Length 5038;  
 Best Local Similarity 22.0%; Pred. No. 0.00019;  
 Matches 254; Conservative 136; Mismatches 435; Indels 328; Gaps 57;

QY 57 RKPA--AGLSAAPVPTAPAAGAPLMDFG---NDFVPPAP -----GPLPAAPPVAPE 103  
 : || || : | | | | : | : | : |  
 Db 288 KSPAQPAGTGKSPAQPPTAKPPAQAGLEKTSLQQPGPKSLAQTPGQGVPPGPAKSPA 347  
 QY 104 RQ-----PSWDPSP-VSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQ-A 155  
 : | : | | : | | | : | | | | : | | | |  
 Db 348 QQPGTAKLPAQQPGPQTASKVPGPTKTPAQLSGPGKTPAQPGPTKPSPPQQPIPAKPQPQ 407  
 QY 156 EPVWT---PPAPAPAAP---STPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQED 209  
 : || | | || | | || | : : : | | | :  
 Db 408 QPVATKPQQPAPAKPQPHPTAKPQPQQPTPA-----KPQPQQPTPAKPQP 456  
 QY 210 FPSVLLETAASXPSLS-----PLSA--ASFKEHEYLGNLSTVLP---TEGTLQENV 256  
 | : : | | | | | | : | : | | |  
 Db 457 QHPGLGKPSAQPSKSIQTVTGRPLQAPPTSAAQAPAQGLSKTICPLCNTTELLHTPE 516  
 QY 7 EASKEVSEKAKTLI-----DRDLTEFSEL-----EYSEMGSSESVSPPKA 296  
 : | : : : : : | | | | : : | : |||  
 Db 517 KANFNTCTECQSTVCSLCGFNPNPHLTEIKEWLCLNCQMQRALGGELAAIPSSPQPTPKA 576



Qy	297	ESAVIV-----ANPREEIIVKN---KDEEEK-----LVSNNILHX-----QQE	331
Db	577	ASVQPATASKSPVPSQQASPKKELPSKQDSPKAPESKKPPPLVKQPTLHGPTPATAPQPP	636
Qy	332	LPTALTCLV--KEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS	389
Db	637	VAEALPKPAPPKPSAALPEQAK-----APV-----ADVEFKQP--KTTETLTDS	679
Qy	390	DMLAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYIT	449
Db	680	PSSAAATSKPAILSSQVQ---AQAVTTAPPLKTDSAKTSQSFPPTGD-----T	725
Qy	450	CAPFN-----PAATESIATNIFLLEDPTSENXTD-EKKIEEKAQIVTEKNTSTKTSN	502
Db	726	ITPLDSKAMPRPASDSKIVSH-----PGPTSESKDPVQKKEEPKKAQTKVTPKPDTK---	777
Qy	503	PFFVAAQDSETDYVTTDNLTKVTEEV-----VANMPEGLTPDLVQEACE	546
Db	778	PVPKGSPTPSGTRPTTGQATPQSQQPPKPEQSRFRSLNLGGIADAPKS-QPTTPQET--	834
Qy	547	SELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSFESEATPSPVLP---DIVM	602
Db	835	-----VTGKLFGFGASI-FSQASNLIS----TAGQQAPHPQTGPAAPSKQAPPPSQTLLA	884
Qy	603	EAPLNSA--VPSAGASVIQ-----PSSSPLEASSVNYESI-----KHEP---ENPP	643
Db	885	QGPPKSTGQHPSAPAKTTAVKKETKGPAAENLEAKPAQAPTVKKA EKDKKHPPGKVS KPP	944
Qy	644	PYEEAMSVSLKVSIGIKEEIKEPENINAALQETEAPYISIACDLIK-ETKLSAEPAPDFSD	702
Db	945	P-----TEPEKAVLAQKPKDKTTKPKPACPLCRTELNVGSQDPPNFNT	986
Qy	703	YSEMAKVE-----QPVPDHSELVE-----DSSPDSEPVD	731
Db	987	CTECKNQVCNLCGFNPTPHLTEIQEWLCLNCQTQRAISGQLGDMDKMPPASSGPKASPVP	1046
Qy	732	LFSD-----DSIPDPVQKQDETIV-----MLVKESLT	757
Db	1047	APAEPPPQKTPTAAHAKGKKKETEVKAETEKQIPEKETPSIEKTPPAVATDQKLESEVT	1106
Qy	758	ETSFESMIEYENKEKLSAL-----PPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSK	809
Db	1107	KSLVSVLPPEKKPSEEEKALPADKKEKKPPAAEAPPLEEKKPIPDQK--LPPDAKPSASE	1164
Qy	810	KEKIPLOMEELSTAVYSNDDLFISKEAQIR-ETETFSDV EIIDEFPTLISSKTPD&FS	868
Db	1165	GEE---KRDLLKAHVQIPEEGPIGKVASLACEGEQQPDTRPEDLPGATPQTLPKD-----	1216
Qy	869	KLAREYTDLEVSHKSEIANAPDGAG--SLPCTELPHDLS--LKNIQPKVEEKISFSDDFS	924
Db	1217	---RQKESRDVTQPQAEGTAKEGRGEPKDRTEKEEDKSDTSSSQPKSPQGLS-DTGYS	1272
Qy	925	KNGSATSKVLLLPPDVSA LGHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSA	984
Db	1273	SDGISGS-----LG-----EIPSLIPSD--EKDLLKGLKKDSFSQESSPS--PS	1313
Qy	985	DLGK--TSVVDLL	995

Db                    || | :|: :|  
1314 DLAKLESTVLSIL 1326

RESULT 13

ANK2\_HUMAN

ID    ANK2\_HUMAN            STANDARD;            PRT;    3924 AA.  
AC    Q01484; Q01485;  
DT    01-APR-1993 (Rel. 25, Created)  
DT    01-OCT-1996 (Rel. 34, Last sequence update)  
DT    28-FEB-2003 (Rel. 41, Last annotation update)  
DE    Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).  
GN    ANK2.  
OS    Homo sapiens (Human).  
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX    NCBI\_TaxID=9606;  
RN    [1]  
RP    SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RC    TISSUE=Brain stem;  
RX    MEDLINE=91302466; PubMed=1830053;  
RA    Otto E., Kunitomo M., McLaughlin T., Bennett V.;  
RT    "Isolation and characterization of cDNAs encoding human brain  
RT    ankyrins reveal a family of alternatively spliced genes.";  
RL    J. Cell Biol. 114:241-253(1991).  
RN    [2]  
RP    REVISIONS.  
RA    Carpenter S.;  
RL    Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.  
RN    [3]  
RP    SEQUENCE FROM N.A. (ISOFORM 1).  
RC    TISSUE=Brain stem;  
RX    MEDLINE=94075409; PubMed=8253844;  
RA    Chan W., Kordeli E., Bennett V.;  
RT    "440-kD ankyrinB: structure of the major developmentally regulated  
RT    domain and selective localization in unmyelinated axons.";  
RL    J. Cell Biol. 123:1463-1473(1993).  
RN    [4]  
RP    SEQUENCE OF 463-495 FROM N.A.  
RX    MEDLINE=92009921; PubMed=1833308;  
RA    Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,  
RA    Lux S.E., Ward D.C., Forget B.G.;  
RT    "Isolation and chromosomal localization of a novel nonerythroid  
RT    ankyrin gene.";  
RL    Genomics 10:858-866(1991).  
CC    -!- FUNCTION: Attach integral membrane proteins to cytoskeletal  
CC    elements. Also bind to cytoskeletal proteins  
CC    -!- ALTERNATIVE PRODUCTS:  
CC    Event=Alternative splicing; Named isoforms=3;  
CC    Name=1;  
CC    IsoId=Q01484-1; Sequence=Displayed;  
CC    Name=2;  
CC    IsoId=Q01484-2; Sequence=VSP\_000267, VSP\_000268;  
CC    Name=3;  
CC    IsoId=Q01484-3; Sequence=VSP\_000268;  
CC    -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL  
CC    CELLS THROUGHOUT THE BRAIN.

CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES  
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE  
CC AND FUNCTION (POTENTIAL).  
CC -!- SIMILARITY: Contains 23 ANK repeats.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to 32Y 00X8 -sib.ch)MID1 094559  
CC -----  
DR EMBL; X56957; CAA40278.1; -.  
DR EMBL; X56958; CAA40279.2; -.  
DR EMBL; Z26634; CAB42644.1; -.  
DR EMBL; M37123; AAA62828.1; -.  
DR PIR; S37431; S37431.  
DR HSSP; P42771; 1DC2.  
DR Genew; HGNC:493; ANK2.  
DR MIM; 106410; -.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 23.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 22.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 20.  
DR PROSITE; PS50297; ANK\_REP\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;  
KW Phosphorylation.  
FT REPEAT 63 92 ANK 1.  
FT REPEAT 96 125 1200XA K 2. VAPAPAX  
FT REPEAT 129 158 ANK 3.  
FT REPEAT 162 191 ANK 4.  
FT REPEAT 193 220 ANK 5.  
FT REPEAT 232 261 ANK 6.  
FT REPEAT 265 294 ANK 7.  
FT REPEAT 298 327 ANK 8.  
FT REPEAT 331 360 ANK 9.  
FT REPEAT 364 393 ANK 10.  
FT REPEAT 397 426 ANK 11.  
FT REPEAT 430 459 ANK 12.  
FT REPEAT 463 492 ANK 13.  
FT REPEAT 496 525 ANK 14.  
FT REPEAT 529 558 ANK 15.  
FT REPEAT 562 591 ANK 16.  
FT REPEAT 595 624 ANK 17.  
FT REPEAT 628 657 ANK 18.  
FT REPEAT 661 690 ANK 19.  
FT REPEAT 694 723 ANK 20.

Feature	Start	End	Score	Description
FT REPEAT	727	756		ANK 21.
FT REPEAT	760	789		ANK 22.
FT REPEAT	793	822		ANK 23.
FT DOMAIN	1773	1950		REPEAT-RICH REGION.
FT REPEAT	1773	1784		REPEAT A.
FT REPEAT	1785	1796		REPEAT A.
FT REPEAT	1797	1808		REPEAT A.
FT REPEAT	1809	1820		REPEAT A.
FT REPEAT	1821	1832		REPEAT A.
FT REPEAT	1833	1844		REPEAT A.
FT REPEAT	1845	1856		REPEAT A.
FT REPEAT	1857	1867		REPEAT A (APPROXIMATE).
FT REPEAT	1868	1879		REPEAT A.
FT REPEAT	1880	1891		REPEAT A.
FT REPEAT	1892	1902		REPEAT A (APPROXIMATE).
FT REPEAT	1903	1914		REPEAT A.
FT REPEAT	1915	1926		REPEAT A.
FT REPEAT	1927	1938		REPEAT A.
FT REPEAT	1939	1950		REPEAT A.
FT DOMAIN	3536			
FT VARSPLIC	1039	1039		Q -> QFLGKLHLPTAPPPLNEGESLVSRIQLGPPGTK (in isoform 2). /FTId=VSP_000267.
FT VARSPLIC	1444	3528		Missing (in isoform 2 and isoform 3). /FTId=VSP_000268.
FT CONFLICT	475	476		GQ -> PE (IN REF. 4).
FT CONFLICT	971	971		I -> S (IN REF. 1).
FT CONFLICT	3581	3582		QY -> HA (IN REF. 1).
FT CONFLICT	3586	3586		I -> Y (IN REF. 1).
SQ SEQUENCE	3924	AA; 430337		MW; 52AC496C428E29D2 CRC64;

Qy	14	DSPP-----RPQPAFKYQFVPEPEDEEEEEEEDEEDEDLE-----EVEVPEAK	589860XT
Db	1648	DIPDETQSTQKQHKPSLGIKKPVRRKLKEKQKQKEEGLQASAEKAEKKGSSSEESLGED	1707
Qy	59	PAAGLSAAPVPTAPAAGAPLMD-----FGNDFVPPAPRG----	92
		:   :     :     :	
Db	1708	P--GLAPEPLPTVKAT-SPLIEETPIGSIKDKVKALQKRVEDEQKGRSKLPIRVKGKEDV	1764
Qy	93	-----PLPAA-PPVAPERQPSWDPS-----VSSTVPAPSPL	123
		:	
Db	1765	PKKTTHRHPAASP SLKSERHAPGSPSPKTERHSTLSSSAKTERHPPVSPSSKTEKHSPV	1824
Qy	124	SAAA-----VSPSKLPEDDEPPA-----RPPPPPPASVSPQAEPVW	159
		:             :       : : :	
Db	1825	SPSAKTERHSPASSSSKTEKHSPVSPSTKTERHSPVSSTKTERHPPVSPSGKTDKRPPV-	1883
Qy	160	TPPAPAPAAPPSTPAAPKRR--GSSGAVVXXXXKIM	TISAGQEDFPSVLLE 216
		:       :   :         :   : :	
Db	1884	SPSGRTEKHPPVSPGRTEKRLPVSPSGRTD-----KHQPVSTAGKTEKHLPVSPSG	1934
Qy	217	TAASXPSLSPLSAAS-FKEHEYLGNLSTVLPT-----EGTLQENVSEASKEVSEKAKT	268
		:       :   :   :     :     :	

Db 1935 KTEKQPPVSPTSSTERIEETMSVRELMKAFQSGQDPSKHKTGLFEHKSQKQKQPQEKQKV 1994

Qy 269 -----LLIDRDLTEFSELEYSEMGSSFSVSPKAES--AVIVANPREEIVKNKDE-- 316  
:| |: : :| : :| :| :| :| :| :|

Db 1995 RVEKEKGPILTQRE-AQKTENQTIKRGQRLPVTGTAEKRGVRVSS----IGVKKEDAAG 2049

Qy 317 -EEKLVSNNI-----LHXQQELPTALTCLVKE-----DEVV 346  
:| :| :| :| :| :| :| :| :| :|

Db 2050 GKEKVLSHKIPEPVQSVPEEESHRESEVPK--EKMADEQGDMDLQISPRKTSTDFSEVI 2107

Qy 347 SSE-KAKDSFNEKRVAVEAPMREEYAD---FKPFERVWEVKDSKED----- 388  
| : | : :| :| :| :| :| :| :|

Db 2108 KQELDNDKYQQFRLSEETEKALHLDQVLTSPFNTPPLDYMKDEFLPALSLQSGALDG 2167

heobS Qy 389 -SDMLAAGGKIESNLESKVDKKCFADSLEQTNHE---KDSESSNDDT FPSTPEGIKDRS 444  
| : | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2168 SSES LN EGVAGSPCGSLMEGTPQISSEESYKHEGLAETPETSPELSF--SPKKSEEQT 2225

Qy 445 GAYITCAPFNPAATESIATNIFLLEDPTSENXTDEKKIEEKKAIQVTEKNTSTKTSNPF 504  
| : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2226 G-----ETKESTKTETTTEIRSEKEHPTTKDITGGS--EERGATVTEDESSTESFQKE 2277

Qy 505 FVAAQDSETDYVTTDNLTKVTEEVVA-NMPEGLTPDLVQACESELNEVTGTKIAYETKM 563  
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2278 ATLGSPKDTSPKRQDDCTGSCSVALAKETPTGLT---EEAACDEGQRTFGSS-AHKT-- 2330

Qy 564 DLVQTSEVMQESLYPAAQLCPSFEESEATPSPVL-----PDI VMEAP--LNSAV 610  
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2331 ---QTDSEAQES-----TATSDETKALPLPEASVKTDTGTESKPQGVIRSPQGLELAL 2380

Qy XA 1 PSAGASVIQ-----PSSSPLEASSVNYESIKHE-PEN--PPPYEEA-MSVSLKVBG 6582  
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2381 PSRDSEVL SAVADDS LAVSHKDSLEASPVLEDNSSHKT PDSLEPSPLKESPCRDSLESSP 2440

Qy 658 IKEEIKE---PEN--INAALQETE--APYISIACDLIKETKLSAEPAPDFSDYSEMAKVE 710  
: : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2441 VEPKMKAGIFPSHFPLPAAVAKTELLTEVASVRSRLLRD-----PDGS--AEDDSLE 2490

Qy 711 QPVPDHSELVEDS-----SPDSEPVDLFSDDSIPDVPQKQDETVMVLKESLTETSFESMI 765  
| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2491 Q-----TSLMESSGKSPLSPDTPSSEEVSYEVT PKTTDVSTPKPAVIHECAEED----- 2539

Qy 766 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKK---EKIPLQMEELS 821  
: || || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2540 DSENGEKKRFTPEE----T EMFKMV---TKIKMF-DELEQEAQKRDYKKEPKQEES 2589

Qy 822 TAVYSNDDL FISKEA-QIRETETFS DSSPIEII DEFPTLISSKTDSF SKLAREYTDLEVS 880  
: | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2590 S---SDPDADCSVDVDEPKHTGSGEDES GV-----PVLVTSES RKVSSSSSES----- 2633

Qy 881 HKSEIANAPDGAGS-----LPCTELPHDL SLKN-----IQPKVEEKISF--SDD 922  
: | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2634 -EPELAQLKKGADSGLLPEPVIRVQPPSPLSSMDSNSSPEEVQFQPVVSKQYTFKMNE 2692

Qy 923 FSKNGSATS KVL LPPDV SALGHTQAEIESIVKPKVLEKEAEKKLPSD-----TEKEDRS 977  
: : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 2693 TQEEP GKSEE-----EKDSESHLAEDRHAVSTEAEDRS 2725

RESULT 14

MAPB\_HUMAN

ID MAPB\_HUMAN STANDARD; PRT; 2468 AA.  
 AC P46821;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain  
 DE LC1].  
 GN MAP1B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=95104835; PubMed=7806212;  
 RA Lien L.L., Feener C., Fischbach N., Kunkel L.M.;  
 RT "Cloning of human microtubule-associated protein 1B and the  
 RT identification of a related gene on chromosome 15.";  
 RL Genomics 22:273-280(1994).

CC -!- FUNCTION: THE FUNCTION OF BR PS IS ESSENTIALLY UNKNOWN.  
 CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES  
 CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST  
 CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS  
 CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN  
 CC STABILIZING MICROTUBULES.  
 CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE  
 CC WITH MAP1A AND MAP1B PROTEINS.  
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC responsible for the binding of MAP1B to microtubules.  
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated  
 CC from MAP1B by proteolytic processing. It is free to associate with  
 CC both MAP1A and MAP1B. It interacts with the amino terminal region  
 CC of MAP1B (By similarity).  
 CC -!- SIMILARITY: TO MAP1A.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L06237; AAA18904.1; -.  
 DR Genew; HGNC:6836; MAP1B.  
 DR MIM; 157129; -.  
 DR GO; GO:0005875; C:microtubule associaEMBL/omplex; TAS.  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 6.  
 KW Microtubules; Repeat; Phosphorylation.

FT CHAIN ? 2468 MAP1 LIGHT CHAIN LC1.  
 FT REPEAT 1878 1894 MAP1B 1.  
 FT REPEAT 1895 1911 MAP1B 2.  
 FT REPEAT 1912 1928 MAP1B 3.  
 FT REPEAT 1929 1945 MAP1B 4.  
 FT REPEAT 1946 1962 MAP1B 5.  
 FT REPEAT 1963 1979 MAP1B 6.  
 FT REPEAT 1997 2013 MAP1B 7.  
 FT REPEAT 2014 2030 MAP1B 8.  
 FT REPEAT 2031 2047 MAP1B 9.  
 FT REPEAT 2048 2064 MAP1B 10.  
 FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY  
 FT KKEE AND KKEI/V REPEATS).  
 SQ SEQUENCE 2468 AA; 270618 MW; 540839CBDF09D461 CRC64;

Query Match 5.0%; Score 296.5; DB 1; Length 2468;  
 Best Local Similarity 21.5%; Pred. No. 0.00041;  
 Matches 264; Conservative 148; Mismatches 459; Indels 355; Gaps 59;

Qy 2 EDLDQSPLVSSS-DSPPRPQPAFKYQ---FVREP-----EDEE 35  
 || :: ||:| || : : | : ::|| | :  
 Db 956 EDGEEHVCVSASKHSPTEDDEESAKAEADAYIREKRESVASGDDRAEEDMDEAIEKGAEQ 1015  
  
 Qy 36 EEEEEEEDEDEDLEEELEV-LERKPAAGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPL 94  
 ||| :|||: || | | : | | | | | : | | |  
 Db 1016 SEEAEDEEDKAEDAREEYEPEKMEAEDYVMAVVDKAAEAGGAEQYGFLLTPTKQLG-- 1073  
  
 Qy 95 PAAPPVAPERQPSWDPSVSPSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQ 154  
 : | | : | | : | | | | : | | | :  
 Db 1074 ---AQSPGREPA--SSIHDETLPGGSESEATA-----SDEENREDQPEEFTATSGYT 1120  
  
 Qy 155 AEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQE-----D 209  
 : | | | | : | : | | | |  
 Db 1121 QSTIEISSEPTPMDEMSTP--- -----RDVMSDETNNETESPSQSESVSVSK 116850XG  
  
 Qy 210 FPSVLLETAASXPS-LSPLSAASFKE-----HEYLGNLSTVLPTEGTLQENVSEAS 259  
 : | | | | : : ||: | | : | : ||: | : : | : :  
 Db 1165 YESSLYSQEYSKPADVTPLN--GFSEGSKTDATDGKDYNASASTISPPSSMEEDKFSRSA 1222  
  
 Qy 260 -----KEVSEKAKTLLIDRDLTEF-----SELEYSEMGS---SFS 291  
 || || | | : | | | : | : | : | :  
 Db 1223 LRDAYCSEV--KASTTLDIKDSISAVSSEKVSPPSKSPSLSPSPSPLEKTPLGERSVNFS 1280  
  
 Qy 292 VSP-----KAESAVIVANPR--EEIIVKN--KDEEEKL-----VSNILHX----- 328  
 : : | | : | : | : : : : | : | : : | :  
 Db 1281 LTPNEIKVSAEAEVAPVSPEVTQEVVEEHCASPEDKTLEVVSPSQSVTGSAGHTPYQSP 1340  
  
 Qy 329 ----QQELPTALTCLVKEDEVVSSEKAKDSFNEKRVAVEAPMREEYADFK-PFERVWEVK 383  
 ||| : : | | : | : : | | | | : : | :  
 Db 1341 TDEKSSHLPTEV--IEKPPAVPVSFEDSDAKDENERASVSPMDEPVPDSESPIEKVLSPL 1398  
  
 Qy 384 DSKEDSDMLAAGGKIESNLESKVDDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDR 443  
 | : : : || | : | | | | : : | : | :  
 Db 1399 RS---PPLIGSESAYESFLSA--DDKASGRGAESPFEKSGKQGSPDQVSPVSE----- 1447  
  
 Qy 444 SGAYITCAPFNPAATESIATNIFLLEDPTSENXTDEKK-----IEKK-----A 488  
 : | | : | : | : || | | : : : : | : | :





OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE OF 1-142 FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;  
 RX MEDLINE=96257242; PubMed=8666295;  
 RA Liu D., Fischer I.;  
 RT "Isolation and sequencing of the 5' end of the rat microtubule-  
 RT associated protein (MAP1B)-encoding cDNA.";  
 RL Gene 172:307-308(1996).  
 RN [2]  
 RP SEQUENCE OF 96-2459 FROM N.A., DOMAIN, AND INDUCTION.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain, and Glial tumor;  
 RX MEDLINE=92347374; PubMed=1639092;  
 RA Zauner W., Kratz J., Staunton J., Feick P., Wiche G.;  
 RT "Identification of two distinct microtubule binding domains on  
 RT recombinant rat MAP 1B.";  
 RL Eur. J. Cell Biol. 57:66-74(1992).  
 RN [3]  
 RP SEQUENCE OF 1541-2459 FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Spinal cord;  
 RX MEDLINE=90059871; PubMed=2555150;  
 RA Rienitz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,  
 RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;  
 RT "Neuraxin, a novel putative structural protein of the rat central  
 RT nervous system that is immunologically related to microtubule-  
 RT associated protein 5.";  
 RL EMBO J. 8:2879-2888(1989).  
 RN [4]  
 RP DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.  
 RX MEDLINE=97405699; PubMed=9260743;  
 RA Ma D., Nothias F., Boyne L.J., Fischer I.;  
 RT "Differential regulation of microtubule-associated protein 1B (MAP1B)  
 RT in rat CNS and PNS during development.";  
 RL J. Neurosci. Res. 49:319-332(1997).  
 CC -!- FUNCTION: The function of brain MAPS is essentially unknown.  
 CC     Phosphorylated MAP1B may play a role in the cytoskeletal changes  
 CC     that accompany neurite extension. Possibly MAP1B Binds to at least  
 CC     two tubulin subunits in the polymer, and this bridging of subunits  
 CC     might be involved in nucleating microtubule polymerization and in  
 CC     stabilizing microtubules.  
 CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate  
 CC     with MAP1A and MAP1B proteins.  
 CC -!- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,  
 CC     cerebellum and cerebrum). Not expressed in liver, spleen, kidney,  
 CC     heart or muscle.  
 CC -!- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic  
 CC     nerve levels are high early in development but decrease during  
 CC     postnatal development and are low in adults. In dorsal root  
 CC     ganglia levels remain high throughout development.  
 CC -!- INDUCTION: By nerve growth factor.  
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC     KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC     responsible for the binding of MAP1B to microtubules.  
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated

CC from MAP1B by proteolytic processing. It is free to associate with  
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region  
 CC of MAP1B (By similarity).  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: TO MAP1A.  
 CC -!- CAUTION: A C-terminal fragment of this protein (residues 1597 to  
 CC 2459) was originally described as neuraxin in Ref.3.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U52950; AAB17068.1; -.  
 DR EMBL; X60370; CAC16162.1; -.  
 DR EMBL; X16623; CAA34620.1; ALT\_SEQ.  
 DR PIR; A56577; A56577.  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 8.  
 KW Microtubules; Repeat; Phosphorylation.  
 FT CHAIN ? 2459 MAP1 LIGHT CHAIN LC1.  
 FT REPEAT 1869 1885 MAP1B 1.  
 FT REPEAT 1886 1902 MAP1B 2.  
 FT REPEAT 1903 1919 MAP1B 3.  
 FT REPEAT 1920 1936 MAP1B 4.  
 FT REPEAT 1937 1953 MAP1B 5.  
 FT REPEAT 1954 1970 MAP1B 6.  
 FT REPEAT 1988 2004 MAP1B 7.  
 FT REPEAT 2005 2021 MAP1B 8.  
 FT REPEAT 2022 2038 MAP1B 9.  
 FT REPEAT 2039 2055 MAP1B 10.  
 FT DOMAIN 559 1035 GLU-RICH.  
 FT DOMAIN 588 786 LYS-RICH (HIGHLY BASIC, CONTAINS MANY  
 FT KKEE AND KKEI/V REPEATS).  
 FT DOMAIN 2224 2312 LYS-RICH.  
 FT CONFLICT 127 127 M -> V (IN REF. 1).  
 FT CONFLICT 140 140 T -> S (IN REF. 1).  
 FT CONFLICT 2112 2112 R -> K (IN REF. 3).  
 FT CONFLICT 2169 2169 L -> I (IN REF. 3).  
 SQ SEQUENCE 2459 AA; 269497 MW; 2E3F6872DEDB8BA2 CRC64;

Query Match 4.9%; Score 289; DB 1; Length 2459;  
 Best Local Similarity 20.9%; Pred. No. 0.0008;  
 Matches 275; Conservative 158; Mismatches 462; Indels 422; Gaps 61;

Qy 30 EPEDEEEEEEEEEDEDEDLEELEVLERKPAAGLSAAPVPTAPAAGAPLMDFGNDFVPPA 89  
 | | ||| |||: :: || :: | | | | : ||:  
 Db 1004 EAEQSEEEGEEEEEDKAEDAREEDHEPDKTEAEDYVMAVVDKAAEAGVTEDQY---DFL--- 1058  
 Qy 90 PRGPLPAAPP--VAPERQPSWDPSVPSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPP 147  
 || | :| |:| | |:| | | | |:| |  
 Db 1059 ---GTPAKQPGVQSPSREPA--SSIHDETLPGGSESEATA-----SDEENREDQPEEF 1106

Qy 148 PASVSPQAEPVWTPPAPAPAAPSTPAAPKRRGSSGAVVXXXXKIMDLKEQPGNTISAGQ 207  
 | : : | | || : | : | | |  
 Db 1107 TATSGYTQSTIEISSEPTPMDEMSTP-----RDVMTDETNNNEETESPSQ 1150  
  
 Qy 208 E-----DFPSVLLETAASXP---SLSPLSAASFKE----HEYLGNLSTVLPTEGTLQENV 255  
 | : | | | : || | : : | : || : :  
 Db 1151 EEFVNITKYESSLYSQEYSKPVVASFNGLSDGSKTDATDGRDYNASASTISPPSSMEEDKF 1210  
  
 Qy 256 SEAS-----KEVSEK-----AKTLLIDRDLTE 277  
 | : : | : || : : || | : | :  
 Db 1211 SKSALRDAYRPEETDVKTGAELDIKDVSDERLSPAKSPSLSPSPSPSPIEKTPLGERSV-N 1269  
  
 Qy 278 FS----ELEYSEMGSFS-VSPKAESAVI---VANPRE---EIIVKKN----- 314  
 || : : | : : || | : | | : : : 29MBF 50XS  
 Db 1270 FSLTPNEIKASAEGEATAVSPGVTQAVVEEHCASPEEKTLEVVSQSVTGSAGHTPY 1329  
  
 Qy 315 ----DEEEKLVSNILHXQQELPTALTCLVKEDEVVSSE--KAKDSFNEKRVAVEAPMRE 368  
 || : : : | : || | : || | || : : : || |  
 Db 1330 QSPTDEKSSHLPTTEVTENAQAVP-----VSFEFTEAKDE-NER--SSISPMDE 1374  
  
 Qy 369 EYADFK-PFERVWE-----VKDSKED---SDMLAAGGKIESNLESKVDKKCFADSL 415  
 | : | | : : : || : | | | : || | | : | : |  
 Db 1375 PVPDESPIEKVLSPLRSPPLIGSESAYEDFLSADDKALGRRSESPFEGKNGKQGFS-- 1432  
  
 Qy 416 EQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNP--AATESIATNIFPLLED- 471  
 | : | | | : : | : | : | : | : | : |  
 Db 1433 -----KESPVS-DLTSPLYQDKQEEKRAGFIPIKEDFSPEKKASDAEIMSSQSALALDE 1485  
  
 Qy 472 -----PTSENXTDEKKIEEKAQIVTEKNTSTKTSNPPFFVAAQDSETDYVTTDNLT 523  
 || : : : | : || | : | : | : | : :  
 Db 1486 RKLGGDGSPTQVDVSQFGSKEDTKMSISEGTVSDKSATPVDEGAEDT---YSHMEGVAS 1542  
  
 Qy 524 VTEEVVA--NMPEGLTPDLVQACESELNEVTGKTIAYETKMDLVQTSEVMQES-LYPAA 580  
 | : || : || | | : : | : : : || | || : : | :  
 Db 1543 VSTASVATSSFPPTTDD-VSPSLHAEVGSFPHSTEVDDSLSVSVVQTPTTFQETEMSPSK 1601  
  
 Qy 581 QLCP-----SFESEATPSPVLPDIVMEAPLNSAV-----P 611  
 : || : : : | : : | : | : |  
 Db 1602 EECPRMSSISPPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSLAMDFSQRQSPDHP 1661  
  
 Qy 612 SAGASVIQ-----PSSSPLEASSVNYESIKHE--PENPPPY-----EEAMSVS-LKVSG 657  
 : || : : | : : | : | | | : || | : : |  
 Db 1662 TVGAGMLHITENGPTVDYSPSDIQDSSLSHKIPPTTEPSYTQDNDLSELISVSQVEASP 1721  
  
 Qy 658 IKEEIKEPENINAALQETE-----APYISIACDLIKE---TKLSAEP----- 696  
 | | : || : | | : : : || :  
 Db 1722 STSSAHTPSQIASPLQEDTLSDVVPFRDMSLYASLASEKVSLEGEKLSPKSDISPLTPR 1781  
  
 Qy 697 -----APDFSDYSEMAK-----VEQPVPDHSELVEDS----- 723  
 : | || : || : | | : : | :  
 Db 1782 ESSPTYSPGFSDDSTSGAKESTAAYQTSSSPPIDAAAAEPYGRSSMLFDTMQHHLALSRL 1841  
  
 Qy 724 -----SPDSEPVDLFSDD---SIPDVP---QKQD 746  
 || | | | : || : :  
 Db 1842 LTTSSVEKDNGGKTPGDFNYAYQKPESTTESPDEEDYDYESHEKTIQAHDVGGYGYEYKTE 1901  
  
 Qy 747 ETVML-----VKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKD 797

Db	1902	RTIKSPCDSGYSYETIEKTTKTPEDGGYSCEITEKTTRTPEEGGYSY-----EISEK	1953
Qy	798	TLLPDEV--TLKKKEKIPLOMEELSTAVYSNDD-----LFISKEA	836
Db	1954	TTRTPEVSGYTYEKTERRRLDDISNGYDDTEDGGHTLGDCSYSYETTEKITSFPESES	2013
Qy	837	QIRETETFSDDSP-----IEIIDEFPTLISSKTDSFSKLAREYT-----DL	877
Db	2014	YSYETTTKTRSPDTSAYCYETMEKITKTPQASTYSYETSDRCYTPERKSPSEARQDVDL	2073
Qy	878	-----EVSH-KSEIANA---PDG---AGSLPCTELPHDLSLKNIQP-----KV	913
Db	2074	CLVSSCEFKHPKTELSPSFINPNLEWFAGEEPTESERPLTQSGGAPPPSGGKQGRQC	2133
Qy	914	EEKISFSDDFSKNKSATSKVLLLPDVSALGHTQAEIESIVKPKVLEKEAEKK-LPSD--	970
Db	2134	DETPPTSVSESAPSQTDSDV---PPETE-----ECPSITADANLDSEDESETIPTDKT	2183
Qy	971	-----TEKEDRSPS-----AIFSADLGKTSVVDLLYWRDIKKTG	1004
Db	2184	VTYKHMDPPPAPMQDRSPSPRHPDVSMVDPEALAEQNLGKALKKDLKEKAKTKKPG	2240

Search completed: January 22, 2004, 16:31:59  
Job time : 25.6482 secs